ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

INFORMATION TECHNOLOGY AND GOVERNMENT DECENTRALIZATION: EXPERIMENTAL EVIDENCE FROM PARAGUAY

ERNESTO DAL BÓ
UC Berkeley

FREDERICO FINAN
UC Berkeley

NICHOLAS Y. LI
CFPB

LAURA SCHECHTER
UW Madison

INFORMATION TECHNOLOGY AND GOVERNMENT DECENTRALIZATION: EXPERIMENTAL EVIDENCE FROM PARAGUAY

ERNESTO DAL BÓ
UC Berkeley

FREDERICO FINAN
UC Berkeley

NICHOLAS Y. LI
CFPB

LAURA SCHECHTER
UW Madison

Standard models of hierarchy assume that agents and middle managers are better informed than principals. We estimate the value of the informational advantage held by supervisors—middle managers—when ministerial leadership—the principal—introduced a new monitoring technology aimed at improving the performance of agricultural extension agents (AEAs) in rural Paraguay. Our approach employs a novel experimental design that elicited treatment-priority rankings from supervisors before randomization of treatment. We find that supervisors have valuable information—they prioritize AEAs who would be more responsive to the monitoring treatment. We develop a model of monitoring under different scales of treatment roll-out and different treatment allocation rules. We semiparametrically estimate marginal treatment effects (MTEs) to demonstrate that the value of information and the benefits to decentralizing treatment decisions depend crucially on the sophistication of the principal and on the scale of roll-out.

KEYWORDS: Decentralization, delegation, bureaucracy, monitoring, marginal treatment effects.

1. INTRODUCTION

IN STANDARD MODELS OF DECENTRALIZATION, principals delegate decision-making authority to agents in order to take advantage of their superior information (Aghion and Tirole (1997), Dessein (2002), Mookherjee (2006), Acemoglu, Aghion, Lelarge, Van Reenen, and Zilibotti (2007), Bloom, Sadun, and Van Reenen (2012)). However, agents may offset some of those gains in pursuit of their own objectives, and principals could po-

https://doi.org/10.3982/ECTA17497

tentially acquire some of the agents' information and improve decision-making without relinquishing control.

A government rolling out a new technology to its front-line service providers faces these trade-offs. In 2014, the government of Paraguay distributed GPS-enabled cell phones to agricultural extension agents (AEAs) and their supervisors. AEAs support farmers by providing, among other things, information about prices and demonstrating best farming practices. The central government hypothesized that the cell phones would allow supervisors to track their AEAs across space and mitigate AEA shirking. As the government had limited resources to pay for phones, questions emerged as to whether supervisors might know which AEAs to treat and which to exclude in order to maximize impact.

We conducted an impact evaluation of the technology, but the question of whether supervisors would assign phones well could not be answered through a standard randomized evaluation. We elicited from supervisors their preference about which AEAs to treat, before assigning phones (still) at random. This novel feature of our experiment allows us to compare treatment effects for AEAs who would have been selected by supervisors in a decentralized roll-out against those who would not have been. If supervisors have and make use of valuable information, supervisor-selected AEAs should respond more to treatment.

Properly assessing the value of decentralized information requires answering two additional questions. First, how much of that information does the principal already have? Second, what is (and what should be) the program scale? If the state will treat everyone, the supervisor's information is not useful. In addition, if the state is unable to differentiate among AEAs and the average cost of treatment exceeds the average benefit, the state should forgo the program altogether. But in this case, decentralized selective targeting may render the program valuable. If the program yields heterogeneous treatment effects, decentralized information (and the program itself) will be valuable only if the treatment response for the recipients is higher than its cost.

To evaluate the value of decentralized information, we estimate models that connect treatment effects to characteristics of AEAs—both those that are observable to us and the state as well as those that are unobservable but drive supervisor selections. Because of our unique design delinking supervisor selection of AEAs and treatment assignment, we can identify the correlation between unobservables and treatment effects. Thus, we decompose the supervisor's informational advantage in terms of observables and unobservables. We then apply the schedule of marginal treatment effects (MTEs) implied by the models to interpolate counterfactual effects under different scales of roll-out. We do this both for the case of supervisor assignment and for centralized approaches that target AEAs based on varying amounts of information. We then determine whether wise use of information on observables can bridge the information gap between supervisors and the centralized state.

Findings. We begin by addressing the standard impact evaluation question: does the new monitoring technology reduce shirking? We find that randomly assigned cell phones have a sizable effect on AEA performance, increasing the share of farmers visited in the previous week by an average of 6 percentage points (pp). This represents a 22% increase over the AEAs in the control group. We find no evidence that treated AEAs increased the number of visits at the cost of conducting shorter ones. Cell phones also improve farmer satisfaction with their AEAs by 0.13 standard deviations.

We then evaluate whether supervisors have valuable information about which AEAs to target. We find that supervisor-selected AEAs respond more to increased monitoring, entirely driving the average increase. Among supervisor-selected AEAs, treatment increased the likelihood that a farmer was visited in the past week by 15.4 pp compared to a

statistically insignificant decrease of 3.6 pp among those who were not selected. This finding corroborates the notion that going down the hierarchy from the top program officers to local supervisors could allow organizations to leverage valuable, dispersed knowledge about how best to allocate resources.[1]

What underlies the informational advantage of supervisors? Having collected a rich dataset on the AEAs that includes information on both cognitive and noncognitive traits, we decompose the value of information into parts reflecting observable and unobservable AEA traits. We estimate our model via a two-step procedure in the spirit of a sample selection approach. We find that commonly observed demographic traits (e.g., sex) and harder-to-measure characteristics such as cognitive ability and personality type do a poor job of explaining supervisors' selection decisions. We conclude that supervisors have valuable information above and beyond observable characteristics of AEAs that a central authority could reasonably collect.

We then apply the model estimates to trace out effects under different roll-out scales and assignment schemes. Several conclusions emerge from our exercise. First, the value of supervisor information is substantial relative to a regime in which the principal simply allocates phones at random. The optimal roll-out under decentralization is 75%, while the largest impact gap between decentralized and random allocation of phones occurs at 53% roll-out. At a roll-out of 75%, supervisor assignment increases farmer visits by 7.7 pp, compared to an increase of 4.8 pp under random assignment. However, prediction exercises based purely on observables allow the state to match or even beat the supervisor selections. Under our most sophisticated centralized approach, in which the principal can target treatment based on an experiment revealing the connection between observable characteristics and response to treatment, the optimal roll-out rate is 70%, raising farmer visits by 9 pp. Because MTEs are negative for some AEAs, the optimal roll-out rate under (nonrandom) centralized and decentralized approaches is below 100%. Maximizing aggregate treatment effects also saves on treatment costs relative to universal coverage.

Related literature. Despite much theoretical progress incorporating the idea that agents hold valuable information, direct empirical evidence of this idea remains elusive. One exception is Duflo, Greenstone, Pande, and Ryan (2018) who find that increased random regulatory scrutiny did not significantly reduce pollution relative to less frequent but discretionary audits by inspectors in India. Our study complements theirs in various ways. Our design experimentally identifies the decentralized counterfactual without relying on functional form assumptions. Also, we allow the appeal of decentralization to depend both on supervisors' informational advantage and on their potential preference biases.

The problem of how best to deploy monitoring technology is similar to the problem of how best to target social programs. In the targeting literature, several studies have shown that local officials rely on local information but may be politically motivated (Bardhan, Mitra, Mookherjee, and Nath (2018)) or use alternative definitions of poverty when defining who is poor (Alderman (2002), Galasso and Ravallion (2005), Alatas, Banerjee, Hanna, Olken, and Tobias (2012)). Those individuals who will benefit most from a particular intervention are not necessarily the poorest, and local officials may want to target individuals with the highest returns (Hussam, Rigol, and Roth (2020), Basurto, Dupas, and Robinson (2020)). Our experiment complements this literature in two substantive ways. By divorcing the estimation of treatment effects from selection into treatment, we are able to show experimentally that local agents have useful information about how best

---

[1]We exploit our experimental design to investigate the possibility of spillovers and of supervisors differentially monitoring selected AEAs, and find no evidence of either effect.

to target an intervention. Moreover, we show how different selection rules can achieve different aggregate treatment effects that affect the calculus of how much to scale up the program and whether to decentralize its implementation.

Our study has parallels to the literature on the use of MTEs to construct policy-relevant counterfactuals (Heckman and Vytlacil (2005)). As in the MTE literature, we express the supervisor's problem of how to prioritize AEAs for treatment as a joint model of potential outcomes and selection as determined by a latent index. Crucially, in our setting treatment is not contingent on selection—AEAs who were randomized into treatment were treated regardless of supervisor selection. Thus, when we compute the MTEs, we do not have to extrapolate to subgroups of "always-takers" and "never-takers" because we only have compliers by design. In this respect, our approach implements a variant of the selective trial designs proposed by Chassang, Padró i Miquel, and Snowberg (2012).[2]

Finally, our study adds to a growing body of experimental evidence on the impact of new monitoring technologies in the public sector. Similar to our setting, some of these studies involve weak or no explicit financial incentives (see Aker and Ksoll (2019) on teacher monitoring and Callen, Gulzar, Hasanain, Khan, and Rezaee (2020) on health inspectors). Other studies overlaid financial incentives affecting front-line providers such as teachers, health workers, police, and tax collectors (Banerjee, Glennerster, and Duflo (2008), Dhaliwal and Hanna (2017), Khan, Khwaja, and Olken (2016)). We contribute to this literature by showing that a new monitoring technology can reduce shirking among agricultural extension agents, despite the absence of associated financial incentives.

2. BACKGROUND

We worked with Dirección de Extensión Agraria (DEAg), the arm of the Paraguay Ministry of Agriculture responsible for overseeing national extension programs. The agency is organized into 19 regional centers, which are in turn composed of 182 municipal Agencias Locales de Asistencia Técnica (ALATs). Each ALAT is comprised of a supervisor and other AEAs. Each AEA works with approximately 80 farmers. Supervisors work with their own farmers and also monitor the other AEAs working in the ALAT. Henceforth, we reserve the label of "AEA" for the nonsupervisory agents who work as extensionists. There are roughly 200 AEAs in DEAg.

Extension services help farmers access services that increase agricultural output, both for personal consumption and sale to market. They provide assistance across six official themes: soil improvement, food security, product diversification, marketing, improving quality of life, and institutional strengthening. To that end, AEAs do not usually offer free goods or services to farmers. Much of what they do is connect farmers with cooperatives, private enterprises, and specialists. AEAs also hold group meetings to lead demonstrations and organize farmer field trips. Nonetheless, most of their daily work involves driving away from the ALAT headquarters in towns to rural areas to make house calls. They use the farm visits to diagnose and address farmer-specific agricultural problems.

In June 2014, the Ministries of Planning and Agriculture coordinated to provide AEAs with GPS-enabled cell phones. This initiative had several objectives. The first was to improve coordination and communication within ALATs. For example, AEAs could use the

---

[2]In that paper, the authors recast randomized control trials into a principal-agent problem. They show theoretically how one can recover the MTEs necessary to forecast alternative treatment assignments by eliciting subjects' willingness to pay for treatment. Instead of eliciting the willingness to pay of AEAs, we elicit the targeting priorities of the supervisor.

phone to photograph diseased crops, circulate photos with their supervisors, and get advice for the farmer from a specialist. But crucially, supervisors could use the phones to monitor where AEAs were at all times, how long they spent in each place, and what they did there (since the AEA was supposed to document all their meetings).

All AEAs already owned their own personal cell phones, but they could use the government-provided phones for work-related calls or messages for free. The phones were not fancy or special, and AEAs were not necessarily keen on getting them. However, they also did not complain when assigned a phone or when their colleagues received one.

3. MODEL

Consider a hierarchy consisting of a principal (i.e., ministerial leadership), a supervisor, and a population of agents (i.e., AEAs). The supervisor is responsible for monitoring the agents. We explore how a new technology alters monitoring, and whether the principal can obtain better results by relying on supervisors to decide which agents to target with the technology.

Agents and monitoring. Each agent i caters to a unit mass of farmers, and chooses an effort level that directly determines the share $s_i \in [0, 1]$ of farmers he visits: $s_i(q_i) = q_i \rho_i + \mu_i$, where $q_i \in [0, 1]$ denotes the level of monitoring, ρ_i denotes the agent's responsiveness to monitoring, and μ_i denotes baseline effort. There is no performance pay—effort s_i is noncontractible.[3] The parameter ρ_i captures both the agent's distaste for being reprimanded when caught not visiting farmers (which tends to make ρ_i positive, raising effort) and any resentment for being monitored (which tends to make ρ_i negative, lowering effort). We allow ρ_i to be negative, as we and de Rochambeau (2020) find to be true for some agents empirically. The share of farmers visited (s_i) must be in the unit interval, but otherwise the joint distribution of ρ_i and μ_i is unrestricted.[4]

New technology and treatment effects. Agent i faces monitoring $q_i = q_l + t_i \Delta q$, where q_l is a status quo level of monitoring (normalized to zero for simplicity), $t_i = 1$ if agent i is treated with the new technology, and zero otherwise, and $\Delta q > 0$ is additional monitoring under treatment. Then effort is

$$s_i = \mu_i + t_i \rho_i \Delta q \equiv \mu_i + t_i \beta_i.$$

If placed under treatment, agent i changes his effort by β_i. We assume that the change in monitoring (Δq) is the same across all AEAs, and all variation in β comes from differences in responsiveness to monitoring (ρ), which is distributed according to $F(\cdot)$.[5] While agents may differ across both μ and ρ, we focus on ρ as the relevant dimension of agent heterogeneity because it drives heterogeneous treatment effects. We refer to different levels of ρ as "types." We denote the roll-out scale of the intervention by m. The Average Treatment Effect (ATE) is $\int_\rho \beta f(\rho) \, d\rho$ where $f(\cdot)$ denotes the marginal distribution

---

[3]In a previous version of the paper (available upon request), we formally derive s_i as the solution to an optimization problem for an agent who desires to avoid being caught shirking, and relate the emerging parameters ρ_i and μ_i to underlying parameters reflecting agent motivation. The model can be extended to consider endogenous monitoring effort by supervisors.

[4]We do not assume any particular correlation between ρ_i and μ_i. The possibility that the state can indirectly learn about who is responsive to monitoring (ρ_i) through who is shirking at baseline (μ_i) is an empirical question, which our design allows us to explore.

[5]This assumption is for simplicity. Alternatively, we could let the variation in β be driven by variation in both ρ and Δq. Our empirical approach allows this.
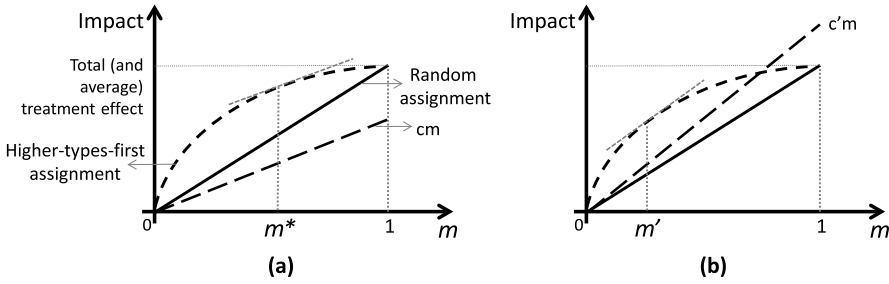
FIGURE 1.—Optimal roll-out and the value of information. The y-axis shows the total treatment effect at different scales of roll-out m depending on whether treatment assignment is random (straight solid line) or ordered by treating more responsive types first (dashed curve—the idealized decentralized allocation). Optimal decentralized roll-out equates MTEs with marginal costs of treatment, and is m^* (m') when marginal cost of treatment is c (c'). Optimal roll-out under random assignment is either 100% when marginal cost of treatment is below ATE (panel a) or 0% when the opposite holds (panel b). In both cases, decentralization attains higher net-of-cost impact, but fully realizing this gain requires knowledge of MTEs.

of ρ. The total treatment impact from targeting a random share m of the population is $m \int_\rho \beta f(\rho)\, d\rho$. The increasing solid line in panel (a) of Figure 1 plots total impact as a function of roll-out scale for a technology assigned at random with a positive ATE.[6]

Three theoretical remarks will guide three distinct blocks of our empirics.

REMARK 1: Suppose $m = 1$ or that agents are selected at random. The total (and average) treatment impact is positive if and only if, given $\Delta q > 0$, the density $f(\cdot)$ places enough weight on positive types.

Thus, a standard impact evaluation comparing average treatment effects between treated and control agents jointly tests that the new monitoring technology increases monitoring intensity ($\Delta q > 0$) and that $f(\cdot)$ places enough weight on positive types. In terms of panel (a) of Figure 1, a typical impact evaluation tests for a positive slope of the impact line under random assignment.

The next two remarks pertain to two stark scenarios capturing centralization and decentralization. Suppose that under *centralization* the principal knows the average effect of the technology, but does not know the underlying agent types. For any given roll-out rate m, the principal can do no better than to select agents at random. In contrast, suppose that under *decentralization* supervisors can observe agent types and prioritize the highest types for treatment.[7] The total impact from treating a measure m of highest types is $\int_\rho \beta \mathbf{1}[\rho > F^{-1}(1 - m)] f(\rho)\, d\rho$, shown by the dashed arc in panel (a) of Figure 1. MTEs, given by the slope of the arc, are initially highest and decrease progressively as the roll-out scale increases and lower types become treated as well. Total treatment impact from supervisor selection and random assignment coincide for $m = 0$ (nobody is treated and impact is zero) and for $m = 1$ (everyone is treated and there is no advantage of prioritizing high types for treatment). Because $E[\beta | \rho > F^{-1}(1 - m)] > E[\beta]$, the remark below naturally follows.

---

[6]Our definition of the total treatment effect abstracts from spillovers across agents. These effects could be modeled, but we do not find evidence of spillover effects as we report below.

[7]In our empirical implementation we let supervisors be less than perfectly informed, and not necessarily benevolent. We will also consider scenarios where the center has some information about agents.

REMARK 2: If supervisors know agents' types and assign treatment to maximize visits to farmers, the average treatment impact on agents selected by supervisors will be higher than on agents selected at random, and on agents who were not selected by supervisors.

The typical concern with decentralization is that supervisors may not select agents according to responsiveness to monitoring (ρ). They may only observe ρ with noise or pursue their own objectives, partly or completely undoing their informational advantage. Nonetheless, they may still outperform random assignment if they have a net informational advantage (i.e., they are not too biased). An experiment like ours that compares the effects of treatment on agents that would have been selected by a supervisor with those who would not have been selected tells us whether supervisors have a net informational advantage.

We stated that the value of decentralization vanishes when m is very high or very low. Decentralization yields value in two ways when choosing roll-out. First, there are programs that would be valuable under centralization, but are even more valuable if decentralization allows to optimize the roll-out scale. Second, decentralization may make programs valuable that would not be valuable under centralization. To illustrate the first case, assume treating an agent costs a constant amount c, so the total cost of the intervention is mc, as shown in panel (a) of Figure 1. Since the ATE exceeds the average cost c, the best centralized scale is 100%. The best decentralized choice attains higher total value at a lower scale m^*, by treating only those types for whom the MTE (weakly) exceeds c. To illustrate the second case, consider panel (b) of Figure 1, where the ATE is smaller than the marginal cost of treatment c'. Under centralized randomized assignment the best roll-out scale is zero—the program is not viable under centralization. But the program is viable at the best decentralized scale of $m' > 0$, attained by treating all types for whom the MTE (weakly) exceeds c'.

REMARK 3: In order to identify the optimal roll-out scale and the value of decentralization, it is necessary to estimate the schedule of marginal treatment effects across agent types.

In the next section, we describe how we estimate the gains from decentralization at all levels of roll-out.

4. RESEARCH DESIGN

Our experiment was conducted in 180 local technical assistance agencies (ALATs, Agencia Local de Asistencia Técnica). All ALATs have a supervisor. There are 46 "large" ALATs that in addition to the supervisor have at least 2 permanent AEAs. On average, these large ALATs have 3.5 AEAs.

Figure 2 presents the research design, with large ALATs on the left. Before conducting the randomization, we asked the supervisor of each large ALAT to indicate which half of her AEAs should receive phones given the program's objective to increase AEA performance.[8] We refer to these AEAs as "selected." AEAs were not told that their supervisor was asked to make such a decision and were not told who was selected. The selected and nonselected are represented by the two rows in Figure 2.

---

[8]The script explained that the ministry was studying whether GPS-enabled cell phones might be a way to strengthen extension services. The script stated that, for budgetary reasons, the ministry would not be able to
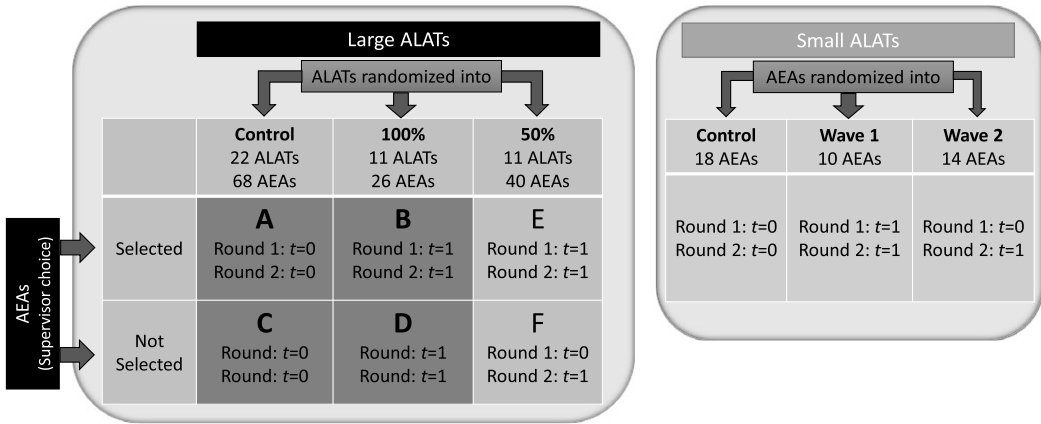
FIGURE 2.—Experimental design. "Wave" corresponds to cell-phone roll-out waves and "Round" to farmer survey rounds conducted after each wave. AEA treatment status (t) is shown in each survey round. Cells A through D constitute our main sample for analysis. The numbers of ALATs and AEAs reflect attrition.

After the supervisors selected AEAs, the large ALATs were randomly divided into three groups, represented by the three columns in Figure 2. The first group of ALATs were assigned to control (cells A and C). There were two waves of phone roll-out (followed by two rounds of phone surveys of farmers) and the AEAs and supervisors in the control condition did not receive phones in either wave. In the second group of ALATs, the 100% group (cells B and D), all AEAs and their supervisors received phones in the first wave. The AEAs in cells A through D constitute the main sample for our analysis. A third group of large ALATs were randomized into the 50% condition. In that group, the selected AEAs and all supervisors received a phone in the first wave (cell E). AEAs who were not selected by a supervisor received a phone in the second wave (cell F).

There are 134 additional "small" ALATs in which no selection by supervisors took place. Of these, 66 have a single permanent AEA in addition to the supervisor. In these ALATs, the AEAs were randomized to receive a phone in wave 1, wave 2, or not at all. The supervisor received the phone in the same wave as the first AEA did.[9] The other 68 ALATs had a supervisor but no AEAs. In these very small ALATs, the supervisors were randomized to receive a phone in wave 1, wave 2, or not at all. Though this randomization occurred jointly, our analysis (other than one robustness check) focuses on the AEAs and excludes these supervisors.

Theory predicts AEA effort to be $s_i = \mu_i + \beta_i t_i$. We allow baseline productivity μ_i and responsiveness to treatment β_i to depend on AEA characteristics X_i and on respective independent residuals ε and η, so $\mu_i = \mu_i(X_i, \varepsilon_i)$ and $\beta_i = \beta_i(X_i, \eta_i)$. Assuming linearity, effort can be written as $s_i = \mu' X_i + \varepsilon_i + (\beta' X_i + \eta_i) t_i$.

---

give phones to all AEAs, so they wanted to know the supervisor's assessment. The script included the statement "It is likely that the ministry will use this information to assign the phones, but for organizational reasons we cannot guarantee that." The representative then asked: "As I mentioned, the goal of giving phones to AEAs is to increase the effort they put into their work as much as possible. Please tell me, which of the AEAs with whom you work do you recommend be selected to receive a phone as soon as possible? Please choose half of the AEAs in your ALAT."

[9]Some small ALATs have more than one AEA if they have AEAs on temporary contracts.

Average treatment effects (ATE). Denote with $E[\beta_i] = \beta_0$ the ATE of cell phones on effort, and with $e_i \equiv \varepsilon_i + (\beta_i - \beta_0)t_i$ the empirical error, to obtain the estimating equation

$$s_i = \mu'X_i + \beta_0 t_i + e_i.$$

The ATE parameter β_0 is identified in our main sample by comparing s_i^* (measured by visits to farmers during the previous week) for AEAs in the 100% treatment ALATs (cells B and D) with the measure for those in control ALATs (cells A and C).[10] Our theoretical Remark 1 stated that $\beta_0 > 0$ if treatment improves monitoring and there are sufficiently many AEAs who respond positively to treatment.

Do supervisors select well? Denote the ATE conditional on selection with $E[\beta_i|D_i] = \beta_0 + \beta_1 D_i$, where the dummy D_i equals 1 for AEAs who are selected by the supervisor. Our estimating equation becomes

$$s_i = \mu'X_i + \beta_0 t_i + \beta_1 D_i t_i + e_i.$$

If supervisor selection leads to superior outcomes compared to random assignment, then β_1 will be positive, aligning with Remark 2 in the theory. The parameter β_1 is identified in our main sample by comparing the difference in visits between selected AEAs in the 100% treatment versus control group (cells B minus A) net of the difference between non-selected AEAs in the 100% treatment versus control group (cells D minus C). Treatment is randomly assigned unconditional on supervisor selection and, therefore, consistent estimates for μ, β_0, and β_1 can be obtained via ordinary least squares.

A framework to decompose supervisor selections and estimate MTEs. We extend our theory to consider supervisors who are neither fully benevolent nor perfectly informed about treatment effects. Denote the value that the supervisor assigns to treating AEA i by

$$v_i = \beta'X_i + \eta_i + \psi'X_i + \zeta_i,$$

which includes treatment responsiveness $\beta_i = \beta'X_i + \eta_i$ but also idiosyncratic (i.e., non-benevolent) supervisor preferences $\psi'X_i + \zeta_i$. The supervisor's preferences reflect characteristics of the AEA that are potentially observable to the analyst (X_i) and others that are not (ζ_i).

We assume the supervisor observes covariates available to us X_i and her own preference term ζ_i, but does not directly observe η_i. Instead, the supervisor observes a signal $\theta_i = \eta_i + \xi_i$, where ξ_i is uncorrelated noise. This implies that the supervisor is only partially informed about the responsiveness of AEAs. Thus, the supervisor maximizes her expected utility by selecting the AEAs that yield her the highest expected value. We can describe this decision rule as supervisors selecting all AEAs above some unobserved threshold, $D_i = 1[E[v_i|X_i, \theta_i, \zeta_i] > c]$. We combine terms and define $\Gamma \equiv \beta + \psi$ and $u_i \equiv E[\eta_i|X_i, \theta_i, \zeta_i] + \zeta_i$. Normalizing the threshold $c = 0$ without loss, we write the supervisor selection rule in its "reduced form" as $D_i = 1[\Gamma'X_i + u > 0]$.

We have modeled the AEA selection criterion of a supervisor who is neither fully benevolent nor perfectly informed. To make further progress, we assume (η, ζ, ξ) are normally distributed. This assumption has two consequences. First, the supervisor selection rule takes the form of a probit regression. Second, the assumptions yield a simple

---

[10]We can expand our sample to include AEAs in small ALATs. AEAs in small ALATs who received a phone in wave 1 are considered treated in both the first and second round of farmer surveys, AEAs who received a phone in wave 2 are considered untreated in the first round and treated in the second round, and AEAs who never received a phone are always considered untreated.

expression for the covariance between unobservable determinants of supervisor decisions and unobservable determinants of treatment effects: $\text{Cov}(u, \eta) = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \sigma_\eta^2$.[11] The structure we have imposed allows us to write $E[\eta_i | X_i, D_i] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \frac{\sigma_\eta^2}{\sigma_u^2} \lambda(X_i, D_i)$ where $\lambda(X_i, D_i) \equiv \sigma_u \frac{\phi(\frac{-\Gamma' X_i}{\sigma_u})}{D_i - \Phi(\frac{-\Gamma' X_i}{\sigma_u})}$.[12] This expression says that the supervisor's expected value of the unobservables driving AEA treatment response can be estimated as a function of the inverse Mills ratio $\lambda(X_i, D_i)$, itself a function of AEA observables and an indicator for supervisor selection.

Substituting for $E[\eta_i | X_i, D_i]$ yields expected AEA performance:

$$E[s_i | X_i, D_i, t_i] = \mu' X_i + \beta' X_i t_i + E[\varepsilon_i | X_i, D_i] + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \frac{\sigma_\eta^2}{\sigma_u^2} \lambda(X_i, D_i) t_i. \tag{1}$$

We estimate Γ using a probit regression and then plug the parameters into the expression for the Mills ratio $\lambda(X_i, D_i)$. We can then use $\lambda(\cdot)$ as a regressor in the specification in equation (1). The coefficient on the interaction of the inverse Mills ratio and treatment is identified because our design disconnects selection from treatment.[13]

If supervisors hold valuable information about the AEAs' responsiveness to treatment based on AEA characteristics that are unobservable to the analyst, then the coefficient on $\lambda(X_i, D_i) t_i$ will be positive and significant. The effect of observable elements is captured by β. From these, we can decompose the supervisor's decision into the contribution of (treatment-response-relevant) observable versus unobservable elements.

Once $\lambda(X_i, D_i)$ and the model in (1) have been estimated, it is possible to obtain predictions of treatment effects for AEAs across the support of observables, distinguishing between the contribution of observables and the (expected) unobservables. We can then estimate MTEs for roll-out rates that cover any desired fraction of AEAs, as required by Remark 3 in the theory section.

5. DATA

We collected two main sources of data. The first is a pen and paper survey of AEAs and supervisors that they filled out independently. The survey contains questions regarding

---

[11]Joint normality yields $E[\eta_i | X_i, \theta_i, \zeta_i] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \theta$, and since $u_i = E[\eta_i | X_i, \theta_i, \zeta_i] + \zeta_i$, independence allows us to write $\text{Cov}(u, \eta) = \text{Cov}(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \theta + \zeta, \eta) = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \sigma_\eta^2$.

[12]The law of iterated expectations allows us to write $E[\eta_i | X_i, D_i] = E[E[\eta_i | X_i, D_i, u_i] | X_i, D_i]$, and the joint normality of (η, u) implies $E[\eta_i | X_i, D_i] = E[\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \frac{\sigma_\eta^2}{\sigma_u^2} u | X_i, D_i] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \frac{\sigma_\eta^2}{\sigma_u^2} E[u_i | X_i, D_i]$. The properties of truncated normal distributions yield $E(u_i | X_i, D_i) = \sigma_u \frac{\phi(\frac{-\Gamma' X_i}{\sigma_u})}{D_i - \Phi(\frac{-\Gamma' X_i}{\sigma_u})} \equiv \lambda(X_i, D_i)$ and we obtain the expression $E[\eta_i | X_i, D_i] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \frac{\sigma_\eta^2}{\sigma_u^2} \lambda(X_i, D_i)$.

[13]Equation (1) shares the same functional form as the "Heckit" selection model. In most settings where the Heckit is applied, the treated are also the selected. In our context, instead, treatment t_i is independently and randomly assigned, and not equal to selection D_i. Thus, we neither have censored potential outcomes nor always-takers and never-takers. We have both a randomized experiment with full compliance and also information about supervisor preferences, allowing us to estimate treatment effects along the full distribution of η_i without relying on extrapolation.

the AEAs' demographics, the digit span test measuring cognitive ability, and the Big-5 personality trait inventory. We combine this inventory into stability and plasticity measures. Stability captures the tendency to be emotionally stable, motivated, and organized and has been found to predict earnings and job attainment. Plasticity is a measure of a person's gregariousness and openness to new experiences. These two meta-traits tend to account for much of the shared variance among the lower order dimensions (DeYoung (2006)). The second source of data is two rounds of farmer phone surveys; one round conducted after each wave of phones was distributed. We called farmers who were beneficiaries of the AEAs and asked questions about their interactions with the AEAs such as how often they saw the AEA and how satisfied they were with their work.

Our main outcomes of interest are these interactions between the AEAs and their beneficiaries. Our pre-analysis plan (written in July of 2014) additionally listed other outcome measures such as information transmission to farmers and farmer profits. We realized early on, prior to the start of the intervention, that it would be infeasible to look for impacts on these outcomes and decided not to collect these data.[14] While our final survey did not contain questions on information transmission or farm profits, we did not think to update the preanalysis plan to update our list of outcomes based on the final survey protocol.

The timeline of events is as follows. In early 2014, the Ministry of Agriculture provided us a list of the names of all the supervisors and AEAs working in the 180 ALATs. From this information, we knew which were the large ALATs (those with one supervisor and at least two permanent AEAs). In March 2014, the supervisors of the 46 large ALATs were contacted to choose which AEAs they would like to prioritize for receiving a phone with the objective of expanding effort in service to farmers. Later in March, we conducted the randomization of ALATs and AEAs into the different treatment groups. In early May 2014, we were given lists of the names and phone numbers of farmers served by each supervisor and AEA. We also then found out which individuals on our original list were no longer employed as AEAs.

The first wave of phones was distributed to the AEAs between May and July, 2014. These phones were given to AEAs in cells B and D (all AEAs in the 100% coverage ALATs) and cell E (selected AEAs in the 50% coverage ALATs) as well as a randomly chosen third of AEAs in the small ALATs. After the first wave of phones was distributed, we conducted the first round of farmer phone surveys from July through September 2014 and the AEA survey during September 2014. We treat AEA characteristics such as sex, age, and the personality indices as being fixed. On the other hand, we treat variables such as the AEAs' perceptions of whether their supervisors know where they are during the working week as potentially being affected by the roll-out of the phones. After completing the first round of surveying, the second wave of phones was distributed in February and March of 2015. These phones were given to the AEAs in cell F (the remaining AEAs who were not selected in the 50% coverage ALATs) as well as another randomly chosen third of AEAs in the small ALATs. We then conducted a second round of farmer phone surveys in April and May of 2015.

---

[14]Several factors dissuaded us from attempting to measure profits and farmer knowledge. First, the farmer phone survey needed to be quick and asking about profits would have taken too long. Second, profits are quite variable and we would not have had power to estimate an impact on them with precision. Third, the timing of the agricultural season, of phone roll-out, and of the farmer surveys did not allow for measurable effects on profits. Finally, Paraguayan agriculture is diverse and conversations with the Ministry of Agriculture did not yield a well-defined piece of information that would be relevant across the entire country to test for effects on farmer knowledge.

The ministry did not give any phones to AEAs who were not on our randomized list. There were a few cases in which phones broke or sick AEAs were not able to pick up their phones. For this reason, we look at intent-to-treat (ITT) estimates using the initial random assignment.

Sample Balance and Attrition.   The original list of names we were given in early 2014 included 268 AEAs and 176 supervisors. There are three sources of attrition from this original list. First, just before the start of the intervention in mid-2014, 219 of the AEAs and 148 of the supervisors were still employed by the ministry. Second, we received lists of farmer names and phone numbers for 195 of the AEAs and 132 of the supervisors still working with the ministry. Finally, we have full survey information on 176 of the AEAs and 117 of the supervisors, which makes up our final sample. The original list of AEAs and supervisors spread across 46 large ALATs (with 180 AEAs) and 134 small ALATs (with 88 AEAs). After attrition, our final sample involves 134 AEAs in large ALATs and 42 AEAs in small ALATs.

Figure A1 gives more detailed information on attrition across the treatment groups for the three stages of attrition. In Table AI, we look at whether attrition of AEAs is correlated with treatment, whether attrition is correlated with AEA characteristics, and whether attrition of AEAs with different characteristics is balanced across the treatment groups. For the full sample, we are limited to looking at those characteristics we know from the electoral rolls for all AEAs: their age, sex, and political affiliation. We do not find any differential attrition across the groups either overall, or of AEAs with different characteristics. There is suggestive evidence that younger AEAs and those with higher plasticity are more likely to attrit, but this is balanced across the treatment groups.

The median AEA in our data listed the phone numbers of 52 farmers with whom he or she worked. Panel A of Table AII does not show any significant imbalance in the number of farmer phone numbers reported by each AEA across the treatment arms in the large ALATs. In the small ALATs, AEAs who received phones in the first wave reported fewer farmer phone numbers.

In each round of farmer phone surveys, we called a random subsample of the farmers listed by the AEAs. We did not call all farmers listed, calling 2599 farmers in the first round and 2606 in the second round for the 176 AEAs who made it to our final sample.[15] Of those, 68% led to completed surveys.[16] Conditional on completing the survey, 70% of farmers confirmed that the AEA that had provided their number worked with them, and thus were asked more detailed questions about their interactions with that AEA.[17] This leads to 2477 phone surveys for AEAs (and 1347 phone surveys for supervisors) in our final sample. Panel B of Table AII shows that attrition from a farmer being on our to-call list to making it to our final sample is balanced across the treatment arms.

Table I presents sample means and a randomization check of the cell phone assignment for various AEA characteristics. The table shows differences between the different treatment groups and the control group for the large and small ALATs separately. Overall, the average AEA age is 38, and 71% of AEAs are male. The AEAs were able to recall an

---

[15]We also called 1294 and 1292 farmers in the first and second rounds for the 117 supervisors in our final sample.

[16]In 18% of cases, we reached voice mail on all five tries, 7% of cases were wrong numbers, 4% were out-of service phone numbers, and 2% of farmers did not agree to complete the survey.

[17]We first asked the farmers to talk about any AEAs with whom they worked. We only mentioned the AEA name we had on record if either the farmer worked with an AEA whose name he couldn't remember or if he did not list the name of the AEA we had on record unprompted.

TABLE I

AEA SUMMARY STATISTICS[a]

	Large ALATs			Small ALATs		
	Ctrl	Diff (T100-C)	Diff (T50-C)	Ctrl	Diff (Wave 1-C)	Diff (Wave 2-C)
Male	0.75	−0.13	−0.27	0.61	0.19	0.10
	[0.44]	(0.13)	(0.095)	[0.50]	(0.17)	(0.17)
Age	37.8	4.39	0.91	36.9	−2.79	2.25
	[10.8]	(2.57)	(3.45)	[11.4]	(4.35)	(4.00)
Married	0.44	0.059	0.084	0.28	0.022	0.079
	[0.50]	(0.087)	(0.098)	[0.46]	(0.18)	(0.16)
Average distance	12.7	−1.36	3.74	10.4	3.27	3.07
	[9.21]	(1.72)	(2.67)	[8.41]	(3.72)	(3.76)
Incumbent party	0.56	−0.020	0.091	0.61	−0.011	0.032
	[0.50]	(0.15)	(0.14)	[0.50]	(0.20)	(0.17)
Digit span	5.19	0.27	−0.22	5.33	0.067	0.52
	[1.07]	(0.25)	(0.19)	[0.84]	(0.41)	(0.35)
Big 5—Stability	−0.057	0.11	0.034	−0.083	0.37	0.33
	[1.13]	(0.24)	(0.20)	[1.13]	(0.44)	(0.32)
Big 5—Plasticity	−0.14	0.33	0.25	−0.50	0.60	0.79
	[1.11]	(0.21)	(0.23)	[1.21]	(0.45)	(0.36)
Num AEAs	68	26	40	18	10	14
Num ALATs	22	11	11	17	10	13

[a]Control group means are reported in columns (1) and (4). Mean differences between control and treatment arms reported in columns (2), (3), (5), and (6). Columns (1)–(3) report statistics for "Large" ALATs and columns (4)–(6) report statistics for "Small" ALATs. Standard deviations are reported in square brackets, and standard errors clustering by ALAT are reported in parentheses.

average of 5.3 digits in the digit span memory test.[18] The average distance between the AEA's office and the farmers he visits is 12 kilometers. Overall, Table I suggest that AEA characteristics are balanced across treatment conditions.[19]

6. RESULTS

Effects of Increased Monitoring on Performance. In Table II, we estimate how much more likely it is that farmers of treated AEAs are visited in the week preceding the phone survey. In columns (1) through (3), our estimation sample includes farmers corresponding to AEAs in both small and large ALATs, excluding those randomized into the 50% treatment ALATs (cells *E* and *F*). In column (1), we present the estimates including only intercepts for small ALATs and survey round. In column (2), we add a set of basic controls (e.g., age and sex), and in column (3), we further augment the specification to include the Big 5 personality meta-traits and digit span. In column (4), we exclude small ALATs in a specification mirroring that of column (3).

We find that monitoring increases farmer visits. Farmers assigned to treated AEAs are six percentage points more likely to have been visited than farmers assigned to untreated AEAs, an increase of 22% over the control group. As expected due to the random assign-

---

[18]For the digit span test, the enumerator read a number out loud and requested the AEA recite it back. The first number was two digits long and the enumerator increased the number of digits until the AEA could no longer recall a number correctly on both of two chances.

[19]We also find the treatment to be balanced across a set of ALAT-level characteristics from the population census, agricultural census, and the 2013 presidential elections. These results are available upon request.

TABLE II

AVERAGE EFFECTS OF RECEIVING A CELL PHONE ON PRODUCTIVITY[a]

	Farmer Was Visited in the Last Week					Supervisor Knows	Log Length of Meeting (Mins)	PCA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.066	0.063	0.063	0.057	−0.0075	0.18	−0.0096	0.13
	(0.031)	(0.026)	(0.026)	(0.030)	(0.035)	(0.10)	(0.059)	(0.071)
Selected	0.011	0.015	0.017	0.011		−0.080	−0.034	−0.073
	(0.033)	(0.027)	(0.029)	(0.029)		(0.14)	(0.036)	(0.065)
Servicer	AEA	AEA	AEA	AEA	Supervisor	AEA	AEA	AEA
Mean of control dep. Var	0.27	0.27	0.27	0.27	0.31	4.59	4.41	−0.042
R^2	0.0050	0.012	0.013	0.012	0.032	0.20	0.012	0.017
Number of phone surveys	1842	1842	1842	1584	1173		1819	1838
Number of AEAs	136	136	136	94	107	126	136	136
Number of ALATs	71	71	71	33	107	65	71	71
Includes small ALATs	✓	✓	✓		✓	✓	✓	✓
Basic Controls		✓	✓	✓	✓	✓	✓	✓
Cognitive controls			✓	✓	✓	✓	✓	✓

[a]Outcomes in columns except (6) are from the farmer phone survey. The outcome in column (6) is from the AEA survey. Columns (1)–(3) and (7)–(8) include farmers assigned to AEAs in large and small ALATs, column (4) limits analysis to farmers assigned to AEAs in large ALATs only, and column (5) limits analysis to farmers assigned to supervisors. All columns exclude the 50% treatment group large ALATs. The outcome in columns (1)–(5) is whether the farmer reported being visited by the AEA in the last week. The outcome in column (6) is an index from 1 ("Strongly Disagree") to 5 ("Strongly Agree") measuring whether the AEA agreed with the statement that their supervisor usually knows where they are during the work week. The outcome in column (8) is the first polychoric principal component from measures of farmer satisfaction, the farmer having received training, and the farmer feeling that the AEA is useful. The number of observations in columns (6)–(8) is slightly lower than that in the first three columns due to missing outcome variables. All regressions include an indicator for survey round and all except column (4) include an indicator for small ALATs. Basic controls include male, age, married, average AEA distance from farmers, and incumbent party. Cognitive controls include digit span, big 5 stability, and big 5 plasticity. Standard errors in parentheses clustered at the ALAT level.

ment of treatment, the estimated impact varies little when we add controls.[20] Results are also similar between large and small ALATs. Overall, the demographic and personality-based controls have little predictive power.[21]

Supervisors are responsible for both supervising the AEAs in their ALAT and serving their own farmers. In column (5), we test the impact of the phone on supervisors' visits to farmers. This column includes all supervisors in our sample other than those in large ALATs in the 50% treatment. We find a small and insignificant impact (point estimate = −0.008; clustered standard error = 0.035). This suggests that the impact of the phone is related to the greater monitoring ability it gives supervisors and not to productivity-enhancing functions of the phone. As a further check that phones help monitor the AEAs, in our survey we asked AEAs whether they agreed that their supervisor usually knows where they are during the work week. We see in column (6) that having a phone significantly increased the extent to which AEAs agreed with that statement.

While treatment led to more visits, this does not by itself imply that the AEAs were more productive or exerted more effort. AEAs could simply be making more but shorter

---

[20]In a previous version of the paper, available upon request, we showed that our findings throughout the paper are robust to using wild cluster bootstrap and randomization inference p-values, which help account for the small number of ALATs under study.

[21]In results not shown here we look at short and long-run impacts of the phones (using the first and second rounds of farmer phone surveys), and find that they are quite similar. The impact of the phones does not diminish over time.

visits. In column (7), we estimate the effect of the phone on the length of visits. The point estimate is not statistically significant and suggests that treated AEAs spent only one percent less time (approximately one minute) on each visit.

In column (8), we investigate effects on other dimensions of AEA performance. We use the first polychoric principal component of three indicators of farmer satisfaction: (1) how satisfied the farmer is with the AEA (1 = not at all, 2 = somewhat, 3 = very); (2) whether the farmer reported that the AEA conducted a helpful training session; (3) whether the farmer reported that the AEA was not helpful at all (indicator is 1 where there is no such report). We find that the additional monitoring improved performance along these dimensions as well: the treatment improved aggregate performance by 0.145 of a standard deviation (clustered standard error = 0.079).

The data we collected does not allow us to directly test whether the treatment affected farmer output. However, the government does have data on farmer production collected by the AEAs prior to the intervention in their "Registry of Assisted Agricultural Families." We merge these data on preintervention production with our data on farmer visits for the control AEAs. When we regress farmer log revenue on having been visited by an AEA, we find that an AEA visit is associated with an increase in log revenue of 0.216 (robust standard error = 0.117). If AEA visits are a statistical surrogate for farmer revenue (Athey, Chetty, Imbens, and Kang (2019)) and there are homogeneous effects of monitoring on revenues across farmers, our experimental estimates suggest that better monitoring may have increased farmer revenue by roughly $0.0656 \times 0.216 = 1.42\%$. Of course, AEAs choose which farmers to visit, which may lead to selection bias. For example, AEAs may visit poorer farmers (biasing the estimate downward) or those who would benefit the most from their visits (biasing the estimate upward). We use the procedure developed in Athey et al. (2019) to bound the magnitude of the bias. Unfortunately, given the sparse data at hand, we estimate bounds that are not particularly informative. Nevertheless, these results do provide suggestive evidence that AEA visits may have had an impact on farmer revenue.

Do Supervisors Have Useful Information?   We investigate the value of the information held by supervisors using a simple difference-in-differences estimator for the large ALATs. We compare the performance of AEAs who were selected and received the phone with the performance of those who were selected but did not receive the phone, net of the difference in performance between those who were not selected and received the phone and those who were not selected and did not receive the phone. Results from Table III show selected AEAs increased the share of farmers visited by approximately 15 pp when treated, accounting for the entire average effect of the phone. This effect reflects a 58% increase relative to selected AEAs in the control group who visited farmers somewhat less frequently than nonselected AEAs.

Spillover and Hawthorne Effects.   The results in Table III suggest that supervisors have useful information about which AEAs would respond most to monitoring treatment. However, because supervisors are also involved in using the technology, several issues remain. First, the theory we laid out in Section 3 suggests a straightforward mapping of treatment effects to aggregate effects by precluding the existence of spillovers that could arise if supervisors are simultaneously reallocating effort. Our design allows for a direct test of this assumption by comparing AEAs in the 50% treatment ALATs to those where all AEAs were either treated or untreated.

In particular, AEAs in control cell C and 50% treatment cell F were both not selected and did not receive phones in the first wave. However, the coworkers of AEAs in cell F

TABLE III

THE INFORMATIONAL ADVANTAGE[a]

	Farmer Visited	
	(1)	(2)
Treated	−0.023	−0.036
	(0.043)	(0.038)
Treated × *Selected*	0.14	0.15
	(0.058)	(0.050)
Selected	−0.033	−0.041
	(0.036)	(0.028)
Mean of control dep. Var	0.27	0.27
R^2	0.0090	0.018
Number of phone surveys	1584	1584
Number of AEAs	94	94
Number of ALATs	33	33
Includes basic controls		✓
Includes cognitive controls		✓

[a]Outcome variable is whether the farmer reported being visited by the AEA in the last week. The analysis sample is restricted to the large ALATs in the control and 100% treatment arms. Both regressions include an indicator for survey round. Basic controls include male, age, married, average AEA distance from farmers, and incumbent party. Cognitive controls include digit span, big 5 stability, and big 5 plasticity. Standard errors in parentheses clustered at the ALAT level.

were treated, allowing us to identify a spillover effect in the first round of phone surveys. Similarly, AEAs in cells B and E were both selected and treated in the first wave, but the coworkers of AEAs in cell E were not treated in the first wave. In column (1) of Table IV, we add the observations from cells E and F to our sample and augment the specification from column (2) in Table III with indicators for cells E and F in wave 1. In the presence of spillovers, AEAs in cell F should respond differently from untreated AEAs in cell C (the former have treated coworkers and the latter do not), and AEAs in cell E should respond differently from treated AEAs in cell B (the former have untreated coworkers and the latter do not). We fail to reject that either of these coefficients are different from zero.

These results suggest that supervisors are not reallocating effort to some AEAs at the expense of others. But two issues still need attention. First, supervisors may put more effort into monitoring the selected and treated AEAs in order to validate their selection to their superiors. In this case, our estimates would not be causal effects of the monitoring technology and instead reflect Hawthorne effects. Second, supervisors may have specifically selected AEAs whom they disliked and, if treated, used the cell phone to harass them. In this case, the estimates can still be interpreted as causal but not as reflective of supervisors having information about permanent AEA traits. We find both of these alternative interpretations unlikely for three reasons.

First, since we do not find evidence of spillovers due to a reallocation of effort, the existence of a Hawthorne effect would imply that supervisors are exerting more effort overall. This would in turn lead the average monitoring effects to be larger in large ALATs than small ALATs where no selection took place and the signaling motive is absent. Column

TABLE IV

SPILLOVERS AND CHANGES IN SUPERVISOR EFFORT[a]

	Farmer Visited		Supervisor Called	Supervisor Met
	(1)	(2)	(3)	(4)
Treated	−0.024	0.058	0.094	0.77
	(0.037)	(0.029)	(0.46)	(0.38)
Treated × *Selected*	0.11		−0.43	−0.57
	(0.048)		(0.68)	(0.49)
Treated × Small		0.030		
		(0.074)		
F1	−0.014			
	(0.049)			
E1	0.053			
	(0.061)			
Selected	−0.047	0.017	0.33	0.43
	(0.038)	(0.029)	(0.47)	(0.39)
Small		−0.011	0.37	−0.58
		(0.046)	(0.40)	(0.35)
Mean of Control dep. Var	0.28	0.27	3.45	4.14
R^2	0.012	0.014	0.11	0.15
Number of phone surveys	2228	1842		
Number of AEAs	134	136	127	127
Number of ALATs	44	71	67	67
Includes small ALATs		✓	✓	✓
Includes Includes 50% Treatment ALATs	✓			
Includes basic and cognitive controls	✓	✓	✓	✓

[a]Outcome variable in columns (1) and (2) is whether the farmer reported being visited by the AEA in the last week. Outcome variables in columns (3) and (4) are from the AEA survey and measure how many times a week the AEA talks on the phone with and meets in person with their supervisor. Column (1) includes observations from all large ALATs. $E1$ and $F1$ are indicators for whether the farmer was called in the first round of the phone survey and works with an AEA in treatment cell E or F. Column (2) adds the small ALATs and excludes the 50% treatment large ALATs. All regressions include an indicator for survey round. Basic controls include male, age, married, average AEA distance from farmers, and incumbent party. Cognitive controls include digit span, big 5 stability, and big 5 plasticity. Standard errors in parentheses clustered at the ALAT level.

(2) of Table IV shows we cannot reject that treatment effects are similar across small and large ALATs. If anything, the treatment effect is slightly larger in the small ALATs.[22]

Second, AEAs were asked how many times per week they spoke with their supervisor by phone or in person. If supervisors were applying differential pressure on the selected AEAs (either to validate their selections or to harass AEAs they do not like), then we would expect that the selected and treated AEAs interact more often with their supervisor. We do not find any evidence that this is the case when using either the number of phone calls in the last week (column (3) in Table IV) or number of live interactions (column (4)). In other words, treatment does not drive the selected to be under any more or less scrutiny.[23]

[22]The point estimates do not change when we drop controls for AEA characteristics, which suggests no systematic differences in AEA traits across small versus large ALATs.

[23]Two additional facts speak against a harassment story. First, our AEA survey elicited the frequency of nonwork-related social interactions between AEAs and supervisors. We find no evidence that supervisors select AEAs with whom they have less frequent social interactions, as one might expect if antipathy drives selection.

Third, if effort effects were present, then we might expect them to be more pronounced in the 50% coverage treatment arm (where supervisors' selection rule was enforced) relative to the 100% coverage treatment. Alternatively, one could conjecture that the effort effects would be less pronounced in the 50% coverage treatment arm if, for instance, supervisors in the 100% coverage treatment felt slighted. These possibilities lack empirical support. Column (1) in Table IV indicates that effects in round 1 are similar for selected AEAs across the two treatment arms.

In sum, we find strong evidence that the phones increase AEA effort and that supervisors possess useful information regarding which AEAs' performance will improve most after receiving a phone. This indicates that in the absence of enough phones to treat all AEAs, or in a setting in which some agents react negatively to treatment, it may be valuable to decentralize to supervisors the assignment of phones. This begs the question of what characteristics the supervisors used to create their prioritized list and whether this is information analysts could hope to obtain. The next subsection answers these questions.

Heterogeneous Treatment Effects.   Columns (1) and (2) of Table V, report a probit to predict selection using observable AEA characteristics. Supervisors selected AEAs who were younger, married, and had to travel longer distances to visit their farmers. Supervisors were more likely to select AEAs with lower levels of the Big-5 Stability meta-trait. Individuals with higher stability scores may be more likely to stay motivated and have better relationships with their supervisors without extra monitoring technologies.

We also find that supervisors in large ALATs, all but one of whom are registered with the incumbent political party, are significantly less likely to place AEAs registered with the incumbent party under increased monitoring. This either suggests that supervisors are acting nonbenevolently, or that party affiliation predicts response to treatment.

Despite our rich data, we cannot predict precisely how supervisors select: the highest pseudo R^2 is only 18.5%. Unobservable determinants may reflect supervisor knowledge of AEA responsiveness (reflecting η) or supervisor preferences (reflecting ζ), which we will explore exploiting our research design.

In columns (3) to (5) of Table V, we present a series of second stage estimates based on equation (1).[24] In column (3), we present a specification without any additional controls or interaction terms, whereas in columns (4) and (5) we include additional controls along with their interactions with the treatment indicator. For columns (4) and (5), the first stage regressions are the ones presented in columns (1) and (2).

Because no controls were included in column (3), the coefficient on the inverse Mills ratio interacted with treatment in that column parallels the findings from Table III that supervisors are selecting individuals who respond more strongly to treatment.

The inverse Mills ratio continues to positively predict responsiveness to treatment (as measured by the coefficient on the interaction of the inverse Mills ratio with the treatment indicator) with the inclusion of AEA observables and their interaction with the inverse Mills ratio. This is direct evidence that the unobservable drivers of supervisor selection are on balance related to treatment responsiveness.

Along observable dimensions, predictors of selection do not seem to correspond to predictors of response to treatment. For example, performing worse on the digit span test

---

Second, if some supervision takes place even without phones, a harassment interpretation would predict that the selected in the control group are under stronger scrutiny and, all else equal, perform better. The coefficients on the *Selected* variable in Table III do not support this prediction.

[24]Similar results (available upon request) obtain when including the 50% coverage ALATs.

TABLE V

SELECTION AND EFFECT HETEROGENEITY ON OBSERVABLES AND UNOBSERVABLES[a]

	AEA Was Selected		Farmer Was Visited in the Last Week		
	(1)	(2)	(3)	(4)	(5)
			Main Effects		
Inverse Mills Ratio			−0.021 (0.023)	−0.019 (0.016)	−0.016 (0.016)
Average Treatment Effect			0.062 (0.035)	0.058 (0.026)	0.064 (0.028)
	Main Effects		Interactions with Treatment		
Inverse Mills Ratio			0.088 (0.036)	0.064 (0.022)	0.064 (0.025)
Male	−0.52 (0.35)	−0.51 (0.40)		−0.021 (0.051)	−0.0039 (0.066)
Age	−0.036 (0.018)	−0.037 (0.017)		−0.0033 (0.0031)	−0.0052 (0.0035)
Married	0.86 (0.46)	0.79 (0.44)		−0.084 (0.061)	−0.094 (0.052)
Average distance to farmers (log)	0.33 (0.17)	0.32 (0.20)		0.10 (0.068)	0.10 (0.077)
Incumbent party	−0.81 (0.36)	−0.84 (0.37)		−0.081 (0.038)	−0.069 (0.048)
Digit span		−0.11 (0.14)			−0.064 (0.027)
Big5—Stability		−0.23 (0.12)			−0.038 (0.049)
Big5—Plasticity		0.12 (0.15)			0.035 (0.025)
(Pseudo)-R^2	0.16	0.18	0.0090	0.023	0.028
p-value for observable interactions				0.031	0.0010
Number of phone surveys			1584	1584	1584
Number of AEAs	94	94	94	94	94
Number of ALATs	33	33	33	33	33
Basic controls				✓	✓
Cognitive controls					✓

[a]Columns (1)–(2) show probit regressions at the AEA level; the dependent variable is whether the AEA was selected by their supervisor. Coefficients on the controls are shown. Columns (3)–(5) show OLS regressions at the farmer level; the dependent variable is whether the farmer reported being visited by the AEA in the last week. Coefficients on the controls interacted with treatment are shown. Uninteracted controls are included but coefficients omitted for space. The inverse Mills ratio in column (3) is computed without any covariates. The inverse Mills ratios in columns (4) and (5) use coefficients from regressions in columns (1) and (2), respectively. The analysis sample is restricted to the large ALATs in the control and 100% treatment arms. Regressions in (3)–(5) include an indicator for survey round. Basic controls include male, age, married, average AEA distance from farmers, and incumbent party. Cognitive controls include digit span, big 5 stability, and big 5 plasticity. Standard errors in parentheses clustered at the ALAT level.

predicts higher response to treatment but is not a strong predictor of being selected. On the other hand, both age and party affiliation are strongly predictive of being selected. However, neither are strongly predictive of response to treatment after the inclusion of cognitive controls. These results imply that if the central government has access to some of these observable characteristics, and knows which observable characteristics are corre-

lated with responsiveness to treatment, then the central government can target treatment despite not having access to the same information as the supervisors.

7. COUNTERFACTUALS

In this section, we apply our heterogeneous treatment effects model to compute counterfactuals under alternative selection rules. This allows us to assess the benefits of decentralization relative to centralization under different informational assumptions.

We define the aggregate benefit of the technology under an arbitrary counterfactual selection rule D_i^{CF}:

$$\Delta Y^{CF} = \int \mathbb{E}\{\beta | D_i^{CF} = 1, X_i\} \Pr\{D_i^{CF} = 1 | X_i\} \, dF_X \qquad (2)$$

As we did with supervisors, we write our arbitrary selection rule as a threshold problem, $D_i^{CF}(X_i, u_i^{CF}) = \mathbf{1}[\mathbf{\Gamma}'^{CF} X_i + u_i^{CF} \geq c^{CF}]$. Because we have not made any distributional assumptions about u_i^{CF}, this does not impose additional assumptions.[25] Combining a selection rule D^{CF} with our treatment effect estimates of $\mathbb{E}\{\beta_i | D_i^{CF} = 1, X_i\}$, we vary the threshold c^{CF} to obtain counterfactual benefits ΔY^{CF} and corresponding roll-out scale $m = \Pr\{D_i^{CF} = 1\}$, the unconditional probability of being chosen.

Panels (a) and (b) of Figure 3 trace out the benefits under different roll-out scales and selection rules. The simplest rule to consider is random allocation, represented by the straight dotted line, which corresponds to an uninformed principal who does not have any information on how to target roll-out. Because selection is unrelated to AEA characteristics, equation (2) reduces to $\Delta Y^{CF} = m \cdot$ ATE.

In contrast, the supervisor's selection rule recovered by the probit outperforms random assignment, represented in Figure 3(a) by the solid curve labeled "Supervisor Preference." The index $\Gamma' X_i$ from the probit regressions in Table V yields $\Pr\{D_i^S = 1 | X_i\} = \Phi(\frac{1}{\sigma_u}(\mathbf{\Gamma}' X_i - c^{CF}))$, and the average treatment effect conditional on being selected is given by $\mathbb{E}\{\beta_i | D_i^S = 1, X_i\} = \beta' X_i + \lambda(D_i^S, X_i, c^{CF})$. Like random assignment, the curve must cross the origin (no impact with no roll-out) and must hit the ATE of 0.064 at full roll-out. The curve must also be consistent with the estimate in Table III, which shows that treatment effects are larger for AEAs selected by their supervisor. The supervisor selection curve crosses 0.070 at 53.8% roll-out, which corresponds to the share of AEAs that received the phones under the actual research design. All other points in the curve are obtained by varying the roll-out scale, which changes the relevant observables X_i and the Mills ratio expression. Thus, the shape of the curve tracing decentralized treatment impact is affected by the variation in observables driving selection and by our normality assumptions.

The supervisor treatment impact peaks at a roll-out of 77% and then declines. Our model admits the possibility that treatment effects are negative ($\rho < 0$), and the result is consistent with the findings in de Rochambeau (2020), who shows that the introduction

---

[25]Note that the assumed cost c^{CF} is not directly observable, and the threshold problem is not a unique representation of the selection rule—any monotonic transformation of the latent index and c^{CF} will yield the same choices. However, we are not trying to directly obtain either of these objects. We only predict the consequences $\Pr\{D_i^{CF} = 1 | X_i, c^{CF}\}$ and $\mathbb{E}\{\beta_i | D_i^{CF} = 1, X_i\}$, which map into the scale of roll-out m and the aggregate counterfactual impact ΔY^{CF}.

(a) Supervisor versus random assignment   (b) Supervisor versus alternative allocation rules
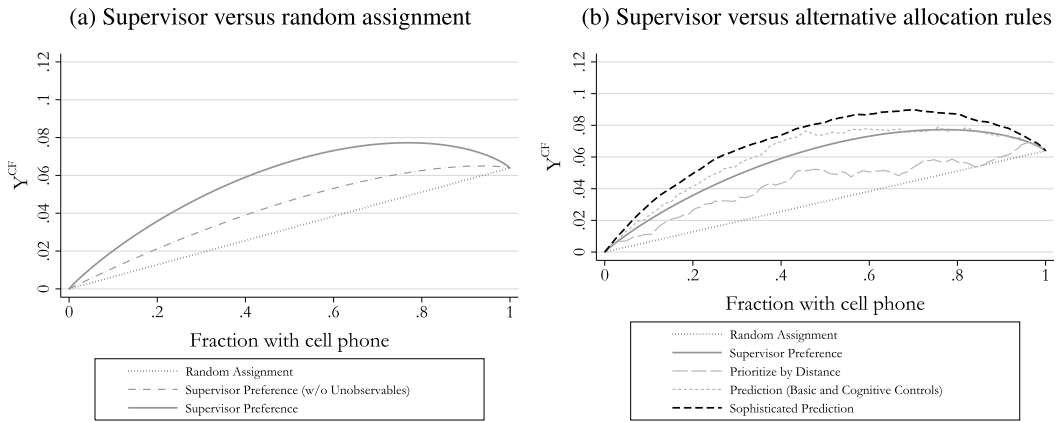

FIGURE 3.—Effects under different allocation rules and roll-out scales. The y-axis shows the total treatment effect at different scales of roll-out under different assignment rules. Under Random Assignment, treatment is assigned to AEAs randomly. Under Supervisor Preference (w/o Unobservables), supervisors select AEAs based purely on observable AEA characteristics. Under Supervisor Preference, supervisors select AEAs based on all the information they have. Under Prioritize by Distance, treatment is assigned first to those AEAs whose beneficiaries live farther from the local ALAT office. Under Prediction (Basic and Cognitive Controls), the observable characteristics of the control group are used to predict baseline performance and AEAs who are predicted to be the worst performers are prioritized. Under Sophisticated Prediction, the principal runs a pilot experiment to establish a map between treatment response and observables, and then prioritizes AEAs predicted to have the highest treatment response.

of a monitoring device for truck drivers in Liberia lowered the productivity of the intrinsically motivated. One advantage of reduced roll-out scales under decentralization is to avoid treating agents for whom treatment backfires.

What underlies the informational advantage of the supervisor? A supervisor has knowledge of characteristics X_i that are observable to the analyst, and of characteristics η_i that are unobservable to the analyst. We assess the contribution of the supervisor's knowledge of unobservables by plotting effects setting $\lambda = 0$, and thus $\mathbb{E}\{\beta_i|D_i^S = 1, X_i\} = \mathbb{E}\{\beta_i|X_i\}$. The effects shown in Figure 3(a) in the dash–dot curve lie above but close to the effects shown in the random assignment line. This suggests that most of the supervisor's informational advantage is driven by access to information that would likely be hard for a centralized authority to collect.

Giving Centralization a Chance. Centralized authorities may not be able to obtain information on η_i. However, they may be able to collect basic information on X_i and use it well. In this section, we consider three alternative counterfactuals varying the level of information available to the central authority.

The first counterfactual corresponds to a *minimally-informed principal* who allocates phones to the AEAs who have to travel the farthest in order to visit their farmers. This requires information on just one observable AEA characteristic. The second counterfactual involves a *partially-informed principal* who can gather information on all AEA characteristics in X_i and map them onto baseline productivity data for a subsample with productivity data. This principal would regress the share of farmers visited in the baseline on AEA characteristics X_i. From this, she would then compute each AEA's expected productivity based on his observable traits. Given this information, the principal would assign cell

phones starting with the AEAs who have the lowest predicted productivity.[26] The third counterfactual assumes a *sophisticated principal* with the capability to conduct a pilot experiment at a low roll-out level in order to establish a mapping between AEA observable characteristics X_i and response to treatment. This principal would regress productivity on characteristics X_i, treatment t_i, and a full set of interactions between them. Then the principal would construct an assignment rule $D_i^{CF}(X_i)$ allocating phones to AEAs predicted to have the highest response to treatment.

We compare these three counterfactuals in Figure 3(b). The minimally-informed principal who only knows distance traveled generally outperforms random assignment (a 2.0 pp advantage at 50% coverage), but she cannot beat the supervisor at any roll-out level. However, the partially-informed and sophisticated principals dominate decentralization at virtually all levels of roll-out.[27] While this conclusion may be context specific, it does suggest that principals who wisely use information that is standard in many employment settings can bridge the information gap.

8. CONCLUSIONS

In this paper, we study an initiative by the federal government in Paraguay to introduce a new monitoring device that enables supervisors to track their agricultural extension agents. We show that the devices induce higher agent effort and that supervisors have valuable information about which agents should be monitored. In addition, we show how the aggregate value of that information varies depending on the share of agents that can be covered with the new technology. We develop an approach to trace out the total treatment effect of the intervention at all levels of roll-out, and compare decentralized assignment against centralized criteria predicated on different data requirements that the central authorities could meet.

Overall our findings suggest that the benefits of decentralization will decrease as information and communication technologies continue to improve. Although studies have shown that innovation in information technologies can lead to more decentralization (Bresnahan, Brynjolfsson, and Hitt (2002), Bloom, Garicano, Sadun, and Van Reenen (2014)), our findings suggest the opposite may occur if new technologies reduce the information gap between principals and agents.

While our specific findings may not be generalizable, our method can be exported to other settings in which spillovers across treatment units are minimal. Future research should extend our framework to incorporate spillovers into the decentralization calculus.

---

[26]If the principal had baseline productivity measures for all AEAs, she could base phone assignment on actual rather than predicted productivity. We cannot estimate this counterfactual because we do not have baseline data from treated AEAs before they received phones.

[27]In these exercises, we use our core experimental sample to both estimate coefficients and plot treatment impact. In reality, the principal may have access to data on a smaller sample than that for which she wants to estimate MTEs, especially in the "sophisticated" case where she may run a small pilot to obtain coefficients and then extrapolate predictions to a much larger set of agents. If this is the case, predictions will incorporate sampling error which we abstract from. An alternative approach would be to split our sample, estimating coefficients in one subsample and predicting MTEs on another. We choose to keep a single sample to preserve power in light of the fact that our main objective is not to compare sampling errors, but to assess whether the information on unobservables that supervisors have confers an unassailable advantage.

TABLE AI

AEA ATTRITION ANALYSIS[a]

	Nonmissing				Nonmissing, Conditional on Employment			
	Large ALATs		Small ALATs		Large ALATs		Small ALATs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat 1	−0.0044	−0.095	−0.060	−0.059	0.0059	−0.58	−0.053	0.20
	(0.085)	(0.40)	(0.14)	(0.68)	(0.061)	(0.41)	(0.15)	(1.15)
Treat 2	−0.0065	0.28	−0.063	−0.35	−0.044	−0.085	−0.053	−0.13
	(0.10)	(0.33)	(0.12)	(0.46)	(0.080)	(0.42)	(0.13)	(0.95)
Male × Treat 1		−0.11		0.23		0.081		0.34
		(0.21)		(0.36)		(0.17)		(0.39)
Male × Treat 2		−0.065		0.17		0.11		0.22
		(0.19)		(0.29)		(0.17)		(0.36)
Age × Treat 1		0.0045		−0.0064		0.0052		−0.0038
		(0.0084)		(0.012)		(0.0047)		(0.024)
Age × Treat 2		−0.0089		0.0044		0.0010		−0.0020
		(0.0091)		(0.010)		(0.0055)		(0.015)
Incumbent Party × Treat 1		−0.12		0.12		−0.14		0.67
		(0.14)		(0.31)		(0.16)		(0.42)
Incumbent Party × Treat 2		0.11		0.00044		0.053		0.45
		(0.18)		(0.25)		(0.17)		(0.39)
Married × Treat 1						0.087		−0.48
						(0.12)		(0.64)
Married × Treat 2						−0.094		0.19
						(0.11)		(0.47)
Digit Span × Treat 1						0.057		−0.094
						(0.045)		(0.18)
Digit Span × Treat 2						−0.0083		−0.0026
						(0.073)		(0.16)
Big 5 Stability × Treat 1						−0.049		0.010
						(0.042)		(0.089)
Big 5 Stability × Treat 2						−0.0083		−0.0062
						(0.030)		(0.082)
Big 5 Plasticity × Treat 1						0.057		0.14
						(0.061)		(0.12)
Big 5 Plasticity × Treat 2						0.051		0.0027
						(0.035)		(0.11)
Basic Controls		✓		✓		✓		✓
Cognitive Controls						✓		✓
p-value of F-test	1.00	0.92	0.85	0.97	0.82	0.62	0.90	0.30
N	180	180	88	88	158	146	61	53

[a]This table tests for differential attrition among the AEAs. In columns (1)–(4), the dependent variable is an indicator for whether an AEA from the original list made it to our final sample. In columns (5)–(8), the dependent variable is an indicator for whether an AEA, who was still employed prior to the intervention, made it to our final sample. The independent variable *Treat 1* corresponds to the 100% treatment arm in the large ALATs and wave 1 in the small ALATs. The independent variable *Treat 2* corresponds to the 50% treatment arm in the large ALATs and wave 2 in the small ALATs. In the even columns uninteracted AEA characteristics are included but coefficients omitted for space. The *F*-test tests the null hypothesis that the treatment effect and (when appropriate) the interaction terms are jointly zero. Basic controls include male, age, and incumbent party. Cognitive controls include married, digit span, big 5 stability, and big 5 plasticity. Standard errors in parentheses clustered at the ALAT level.

TABLE AII

PHONE ATTRITION ANALYSIS[a]

	Large ALATs				Small ALATs			
	N	Ctrl	T100-Ctrl	T50-Ctrl	N	Ctrl	Wave 1-Ctrl	Wave 2-Ctrl
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Phone numbers per AEA	134	49.9	4.80	4.93	42	58.3	−18.5	2.67
			(7.08)	(7.25)			(9.18)	(10.1)
Panel B:								
Phone number in final sample	4573	0.49	0.037	−0.038	632	0.41	0.012	−0.0085
			(0.033)	(0.045)			(0.10)	(0.060)

[a]Panel A shows AEA-level data. It compares the number of farmer phone numbers reported by each AEA across the treatment arms. Panel B shows farmer-level data. It compares the share of farmer phone numbers that made it to our final sample across the treatment arms. Standard errors in parentheses clustered at the ALAT level.
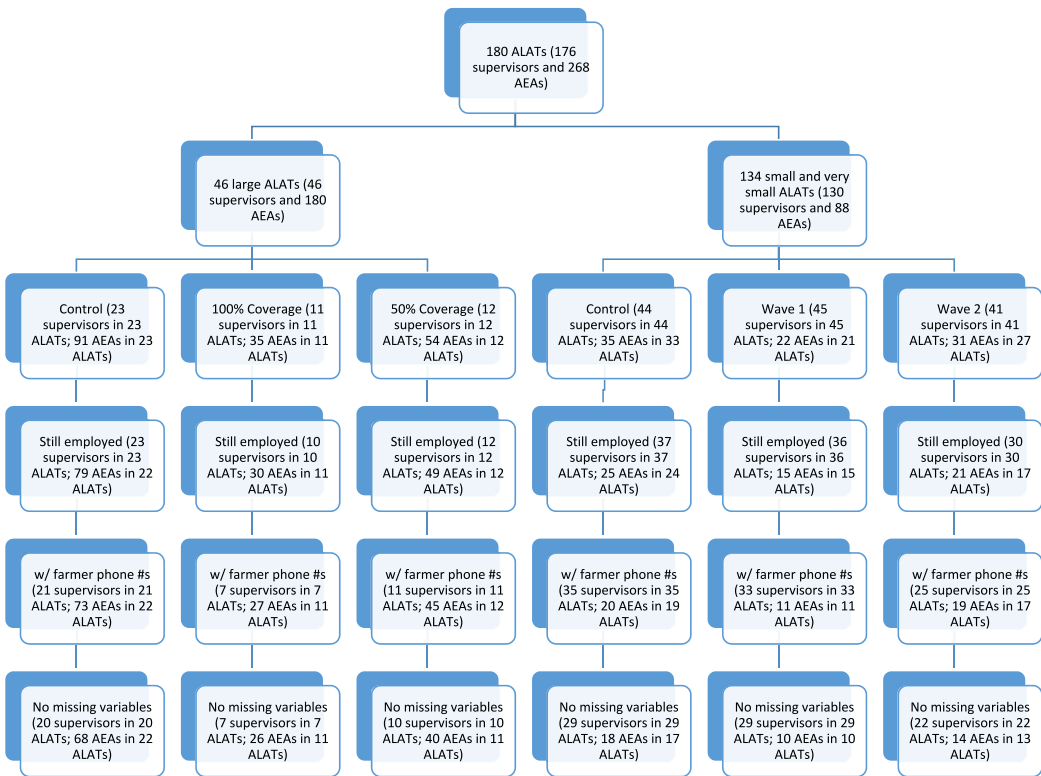

FIGURE A1.—Study design. Columns depict how AEAs and ALATS were allocated across treatment arms. Rows correspond to different stages of attrition.

REFERENCES

ACEMOGLU, D., P. AGHION, C. LELARGE, J. VAN REENEN, AND F. ZILIBOTTI (2007): "Technology, Information, and the Decentralization of the Firm," *Quarterly Journal of Economics*, 122 (4), 1759–1799. [677]

AGHION, P., AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105 (1), 1–29. [677]

AKER, J. C., AND C. KSOLL (2019): "Call Me Educated: Evidence From a Mobile Phone Experiment in Niger," *Economics of Education Review*, 72, 239–257. [680]

ALATAS, V., A. BANERJEE, R. HANNA, B. A. OLKEN, AND J. TOBIAS (2012): "Targeting the Poor: Evidence From a Field Experiment in Indonesia," *American Economic Review*, 102 (4), 1206–1240. [679]

ALDERMAN, H. (2002): "Do Local Officials Know Something We Don't? Decentralization of Targeted Transfers in Albania," *Journal of Public Economics*, 83 (3), 375–404. [679]

ATHEY, S., R. CHETTY, G. W. IMBENS, AND H. KANG (2019): "The Surrogate Index: Combining Short-Term Proxies to Estimate Long-Term Treatment Effects More Rapidly and Precisely," NBER Working Paper 26463. [691]

BANERJEE, A., R. GLENNERSTER, AND E. DUFLO (2008): "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System," *Journal of the European Economic Association*, 6 (2-3), 487–500. [680]

BARDHAN, P., S. MITRA, D. MOOKHERJEE, AND A. NATH (2018): "Resource Transfers to Local Governments: Political Manipulation and Household Responses in West Bengal," Unpublished Working Paper. [679]

BASURTO, M. P., P. DUPAS, AND J. ROBINSON (2020): "Decentralization and Efficiency of Subsidy Targeting: Evidence From Chiefs in Rural Malawi," *Journal of Public Economics*, 185. [679]

BLOOM, N., L. GARICANO, R. SADUN, AND J. VAN REENEN (2014): "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 60 (12), 2859–2885. [698]

BLOOM, N., R. SADUN, AND J. VAN REENEN (2012): "The Organization of Firms Across Countries," *Quarterly Journal of Economics*, 127 (4), 1663–1705. [677]

BRESNAHAN, T., E. BRYNJOLFSSON, AND L. M. HITT (2002): "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics*, 117 (1), 339–376. [698]

CALLEN, M., S. GULZAR, A. HASANAIN, M. Y. KHAN, AND A. REZAEE (2020): "Data and Policy Decisions: Experimental Evidence From Pakistan," *Journal of Development Economics*, 146. [680]

CHASSANG, S., G. PADRÓ I MIQUEL, AND E. SNOWBERG (2012): "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments," *American Economic Review*, 102 (4), 1279–1309. [680]

DAL BÓ, E., F. FINAN N. Y. LI, AND L. SCHECHTER (2021): "Supplement to 'Information Technology and Government Decentralization: Experimental Evidence From Paraguay'," *Econometrica Supplemental Material*, 89, https://doi.org/10.3982/ECTA17497. [677]

DE ROCHAMBEAU, G. (2020): "Monitoring and Intrinsic Motivation: Evidence From Liberia's Trucking Firms," Unpublished Manuscript. [681,696]

DESSEIN, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69 (4), 811–838. [677]

DEYOUNG, C. G. (2006): "Higher-Order Factors of the Big Five in a Multi-Informant Sample," *Journal of Personality and Social Psychology*, 91 (6), 1138–1151. [687]

DHALIWAL, I., AND R. HANNA (2017): "The Devil Is in the Details: The Successes and Limitations of Bureaucratic Reform in India," *Journal of Development Economics*, 124, 1–21. [680]

DUFLO, E., M. GREENSTONE, R. PANDE, AND N. RYAN (2018): "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India," *Econometrica*, 86 (6), 2123–2160. [679]

GALASSO, E., AND M. RAVALLION (2005): "Decentralized Targeting of an Antipoverty Program," *Journal of Public Economics*, 89 (4), 705–727. [679]

HECKMAN, J. J., AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73 (3), 669–738. [680]

HUSSAM, R., N. RIGOL, AND B. ROTH (2020): "Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field," Unpublished Manuscript. [679]

KHAN, A. Q., A. I. KHWAJA, AND B. A. OLKEN (2016): "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 131 (1), 219–271. [680]

MOOKHERJEE, D. (2006): "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective," *Journal of Economic Literature*, 44 (2), 367–390. [677]