Uncertainty, self-selection and the design of subsidies: Evidence from Zambia*

Samuel Bell
Cornell University

B. Kelsey Jack
Tufts University and NBER

Paulina Oliva
UCSB and NBER

Chris Severen
UCSB

Elizabeth Walker
Harvard Kennedy School

Preliminary and Incomplete
Do not Circulate, Do not Cite

Abstract

In the absence of uncertainty, economic incentives for the adoption of a technology with positive environmental externalities can achieve the same adoption rates whether they are conditioned on the acquisition of inputs (take up) or on outcomes. If, on the other hand, potential adopters face uncertainty in the costs and benefits of implementation, conditioning incentives on outcomes may lead to both a different composition of adopters and level of adoption than will subsidizing inputs. We use a field experiment to investigate the correspondence between take up decisions and outcomes for a carbon-mitigating agricultural technology, and observe results consistent with the presence of uncertainty about implementation costs. The experimental variation identifies a structural model of intertemporal decision making under uncertainty that explains the experimental findings and informs welfare estimates. We further use the estimated model to study how different incentive structures perform as we vary assumptions about monitoring costs.

*Helpful comments were received from Jim Berry, Andrew Foster, Andrew Plantinga, Stephen Ryan, Kenneth Train and seminar and workshop participants at Harvard University, CU Boulder, Cornell University, Resources for the Future, UC Berkeley, UC Davis, UC Santa Barbara, SITE 2013, the 2012 AERE meetings, and NEUDC. The authors gratefully acknowledge financial support from the International Growth Centre, the Climate and Development Knowledge Network and Musika. Field work was facilitated by Innovations for Poverty Action, with specific thanks to Jonathan Green, Farinoz Daneshpay, Mwela Namonje and Monica Banda. The project would not have been possible without the collaboration and support of Shared Values Africa and Dunavant Cotton, Ltd.

1 Introduction

Subsidies to increase adoption of technologies with socially desirable outcomes are widespread policy instruments. Whether the subsidy is conditioned on take up (subsidy on purchase price) or on use of the technology (a performance reward) should not affect outcomes, as long as take-up and use decisions are made under the same information sets, i.e. the adopter faces no uncertainty about the benefits and costs associated with use. In the absence of uncertainty, subsidizing inputs may be the more cost effective type of subsidy because it avoids the administrative costs of monitoring outcomes.

If, however, the link between take up and use decisions is broken by the presence of uncertainty, the marginal buyer may differ from the marginal user, and a subsidy conditioned on take up is not equivalent to a performance reward.[1] A potential adopter who faces uncertainty in the cost of using the technology may take up if the expected net benefits are positive but, in the case of a negative cost shock, may prefer not to use the technology. Thus, subsidizing inputs and use no longer guarantees the same provision of outcomes.

The present study uses a field experiment to investigate whether the correspondence between farmer decisions to take up and to use a carbon-reducing agricultural technology are consistent with the presence of uncertainty about costs of implementation.

More specifically, we implement a tree-planting program in Zambia that introduces random variation in subsidies and performance rewards to encourage the adoption of trees that both fix nitrogen (private fertilizer benefits) and sequester carbon (generate an externality in the form of climate change benefits). Our reduced form results show that farmers are responsive to both types of instruments, but on different decision margins. Our reduced form results are consistent with the presence of uncertainty in implementation costs: a large share of farmers take up but do not use the technology (keep trees alive), and we find no difference in performance across levels of the input subsidy. In other words, farmers appear to have limited ability to predict their costs of implementation at the time of take up and self-selection into the program does not correlate well with performance rates. These findings are consistent with a high degree of uncertainty about the expected profits associated with adoption at the time of take up.

[1] Subsidies on take up and performance rewards may also deliver different outcomes if use has an intensive margin. This could be the case of durable goods such as bed nets, light bulbs, solar water heaters, stoves, etc. (Dupas 2013; Ashraf et al. 2010; Cohen and Dupas 2010; Berry et al. 2012) However, to the extent that take up and use decisions are made at different points in time, uncertainty about benefits and costs associated with use may drive an additional wedge between subsidies and performance rewards in the case of durables.

We then present a model of technology adoption under uncertainty.[2] The model allows for some share of the costs associated with implementation to be observed at the time take up decisions are made, and the remaining share to be observed at the time the implementation decision is made. If the share of unobserved costs is large, farmers' participation decisions may be poorly correlated with implementation. In other words, while subsidies may increase take up rates, implementation rates may not improve as quickly, lowering the cost effectiveness of the subsidy. In this case, a reward on outcomes has the advantage of paying out only when the socially desirable outcome is realized. However, note that the reward may increase take up among farmers with all levels of observed implementation costs by insuring against negative shocks to the outside option. In other words, the reward increases the option value of the contract: farmers can abandon the technology without penalty, which implies that the reward only affects the upside of taking up. This increase in take up particularly hurts the cost effectiveness of performance rewards when monitoring costs are high.

Whether it is optimal to subsidize inputs, reward outcomes or a combination of the two depends on these tradeoffs between participation and the delivery of outcomes, as well as the cost of monitoring outcomes. It is also worth noting that a comprehensive welfare calculation should account for farmer's benefits from the program, some of which will come in the form of option value. Hence, whenever evaluating the optimal strategy from a social planner's perspective, the type of instrument chosen may play a role in total net benefits from the program and not only in the distribution of benefits between principal and agent.

The intertemporal decision model we propose can be estimated using the variation created by our experiment. First, note that the heterogeneity in the cost component that is always observed to the farmer creates heterogeneity in expected benefits across farmers at the time of take up. Although some of this heterogeneity may be linked to observable characteristics of the farmer, the model allows for a random component of observed costs, which is identified through the variation in the subsidy for take up. Second, the remaining component of the implementation costs is only observed by the farmer at the time the implementation decision is made. This component is also treated as random (unobserved to the econometrician) and its distribution is identified from farmers' response to the random variation in the rewards for performance.

We use the model estimates to evaluate the role that uncertainty plays in the relative

[2]The model assumes that liquidity constraints are not correlated with the net benefits of the technology, which would confound willingness to pay for the technology with ability to pay. We rule out credit constraints in our empirical setting through our study design, allowing us to focus on valuation of the technology independent on liquidity or credit confounds.

cost-effectiveness of subsidies and rewards as policy instruments. First, we examine the effect of decreasing the uncertainty on the screening power of subsidies and the option value of the rewards, as well as on participation rates. Second, we assess the optimal combination of subsidies and rewards. Third, we examine the role that monitoring costs have in the relative ranking of these policy instruments for a given level of uncertainty.

Our study relates to a number of well developed literatures. First, individual heterogeneity and self-selection in technology adoption are increasingly important themes in development economics (see recent papers by Suri 2011; Chassang et al. 2012; Berry et al. 2012).[3] Much of the existing literature has focused on the difference between the take up margin and the usage margin in durable goods. This difference creates a trade off between the screening effect of prices, which can target products to the households most likely to use the product (Ashraf et al. 2010), and the concern that prices exclude the poorest and potential most vulnerable households (Cohen and Dupas 2010).[4] Like Dupas (2013), we are interested in the dynamics of this tradeoff, but in a setting where the policy maker has two tools at her disposal: subsidies for take up and payments conditional on outcomes. In addition, we offer a conceptual framework that allows for the gap between take up and use margins to emerge from heterogeneity in private benefits across farmers as well as shocks to the opportunity cost of follow through.

Second, we estimate a model of intertemporal decision making under uncertainty, which has been studied both in the theoretical and empirical literature in numerous and varied applications (Arrow and Fisher 1974; Dixit and Pindyck 1994; Rust 1987). Like the Arrow-Fisher and Dixit-Pindyck frameworks, we consider the dynamic value of information, where the initial take up decision in our setting is analogous to investment in period 1 and the decision to deliver outcomes is analogous to investment in period 2. Most relevant to the problem we study are the handful of papers that consider land use change under uncertainty, many of which are applied to data from the Conservation Reserve Program in the United States (Schatzki 2003; Isik and Yang 2004). These applications are similar to the Arrow-Fisher set up in that the option value is associated with delayed investment. Farmers in our study secure the option value of the contract by participating in the program. The take

[3]The importance of unobservable heterogeneity is not, of course, limited to development contexts. Unobservable individual level heterogeneity affects the performance of markets ranging from retirement accounts (Einav et al. 2010) to consumer credit (Karlan and Zinman 2009) to health (Einav et al. 2013).

[4]Prices are not the only approach to improving targeting. Subsidy and social transfer programs often incorporate design elements to induce self-targeting on unobservable traits correlated with socially desirable outcomes (Besley and Coate 1992; Alatas et al. 2012; Dupas et al. 2013).

up action grants them the potential to comply and receive the reward.[5] As in the Arrow-Fisher and Dixit-Pindyck frameworks, the option value of the contracts we offer increases with uncertainty in the net benefits. From a social planner perspective, this value should be incorporated in the welfare calculations when evaluating the cost-effectiveness of the contracts.

Finally, our application, land use subsidies to address environmental externalities, is of increasing policy importance. This paper contributes to the growing literature on the design of payments for environmental services and carbon offsets. A number of descriptive studies analyze the performance of land use payments in developing countries (Alix-Garcia et al. 2012; Pfaff et al. 2008; Jayachandran 2013) or describe the issues and challenges facing contract design (for example, Cacho et al. 2003). Analytical or simulation-based studies of optimal contract design suggest a number of creative ideas for overcoming the challenges of uncertainty, weak enforcement, limited liability and asymmetric information (MacKenzie et al. 2012; Mason and Plantinga 2013; Cordero Salas et al. 2013). However, few papers use exogenous variation to empirically test subsidy design (Pattanayak et al. 2010). An exception is Jack (2013), who demonstrates efficiency gains from targeting on private information about the opportunity cost of afforestation in a PES program in Malawi. In contrast with this earlier work, we find no evidence of self-targeting in our setting, which may be due to differences in how contracts were allocated or in the magnitude of uncertainty in implementation costs. By documenting the negative effect of uncertainty on the social value of environmental land use subsidies, we provide a micro-foundation for Conte and Kotchen's (2010) finding that forestry projects in developing countries are substantially discounted on the voluntary carbon market.

The rest of the paper proceeds as follows. We turn next to a description of the empirical context and experimental design. In Section 3, we show the reduced form experimental results, which inform the model of intertemporal decision making under uncertainty. We present the model and its identification in Section 4 and show estimation results and welfare simulations in Section 5. Section 6 presents alternative explanations, including learning and procrastination, and we conclude with a discussion of the policy relevance of the findings.

[5] In this way, the option value we study is most similar to what is described in the education literature associated with the decision to enroll in college (e.g., Stange 2013).

2 Context and experimental design

The experiment took place in Zambia, where nearly three-quarters of the rural population is engaged in agriculture (CSO 2010), and the potential for afforestation and reforestation is high. Previous studies offer a conservative estimate of the potential for over a billion tons of carbon dioxide equivalent in Zambia from 74,000 square kilometers of land eligible for carbon offsets (Zomer et al. 2008). At a conservative price of USD 5 per ton, the total potential revenue from carbon offsets is upward of USD 5 billion. The project encourages the adoption of agroforestry trees (*Faidherbia albida*) by small holder farmers engaged in contract farming arrangements with a large cotton company. The study was implemented in coordination with Dunavant Cotton Ltd., a large cotton growing company with over 60,000 outgrower farmers in Zambia, and with an NGO, Shared Value Africa. The project, based in Chipata, Zambia, targeted approximately 1,300 Dunavant farmers, organized into 125 outgrower farmer groups. The NGO partner is using the project as an entry point into carbon market development in Zambia, while the private sector partner hopes to improve soil fertility and long run agricultural yields for its farmers.

We describe the study sample, the experimental design and implementation in the remainder of the section.

2.1 The technology

Faidherbia albida is an agroforestry species endemic to Zambia that fixes nitrogen in its roots and leaves, providing farmers with private benefits over the long term, including better soil fertility and higher maize yields. Optimal spacing of *Faidherbia* is around 100 trees per hectare, or at intervals of 10 meters. The relatively wide spacing, together with the fact that the tree sheds its leaves at the onset of the rainy season, means that incorporation of *Faidherbia* on a plot does not displace other crop production (Akinnifesi et al. 2010). Agronomic studies suggest yield gains from planting *Faidherbia* relative to production without fertilizer, with some estimates of increases of 100 to 400 percent (Saka et al. 1994; Barnes and Fagg 2003). However, these private benefits take 7-10 years to reach their full value, and may be insufficient to justify the up-front investment costs in the absence of more short-run incentives. Indeed, a back of the envelope estimate suggests that the adoption of *Faidherbia* is not a privately profitable investment for most farmers, which average benefits of around

USD 3.31 and average costs of around USD 19. [6]This is consistent with low adoption rates the baseline: less than 10 percent of the study households reported any *Faidherbia* on their land.[7]

Although the tree offers numerous potential public benefits, including erosion control and wind breaks, in our analysis we focus on the carbon sequestration benefits. Sequestration flows are estimated from allometric equations based on Brown (1997) and adapted to the growth curves for *Faidherbia*. Figure 1 shows both the annual and cumulative CO_2 take up per tree. Over the course of 30 years, a tree sequesters around 4 tons of carbon. Discounting the sequestration at 0.15 leads to a present value of around 0.48 tons. At a marginal cost of carbon of 20 USD/ton, the sequestration associated with 35 trees is around USD 336. This, of course, assumes no uncertainty.

2.2 Study sample

Project implementation relied on Dunavant's outgrower infrastructure, which is organized around sheds, each of which contains several dozen farmer groups. The farmer groups consist of 10-15 farmers and a lead farmer, who is trained by Dunavant each year and in turn trains his or her own farmers on a variety of agricultural practices. The trainings are typically held in early September through November, immediately preceding planting. The experiment was implemented across six of the Dunavant sheds and reached 125 of the over 200 farmer groups associated with these sheds.[8] Lead farmers self-selected into the training, however, treatment assignment occurred after the study sample was determined, and is therefore orthogonal to lead farmer self-selection.

[6]Agronomic studies suggest yield gains from planting *Faidherbia* relative to production without fertilizer, with some estimates of increases of 100 to 400 percent (Saka et al. (1994); Barnes and Fagg (2003)). The economic value of this increase is around USD 400 per hectare per year. We take mid-range values from the literature and assume that yields on half a hectare of land increase by 525 kilograms, starting 10 years after trees are planted. At current maize prices, this equals additional income of 136.50 USD per year from years 10 to 20. At a discount rate of 0.6, which is what we use in our estimation of the structural model, the present value of this investment is USD 3.31. While these discount rates are high, they ignore risk and uncertainty and are not implausible in rural developing country settings (Conning and Udry (2007)). With respect to the private costs, surveys implemented in conjunction with our study indicate that farmers invest around 38 hours planting and caring for the trees. With an agricultural wage rate of around USD 2.5 per day, these time investments result in around USD 19 of labor costs incurred in the first year.

[7]Informal land tenure in the project area presents an additional barrier to adoption. By focusing on landholders engaged in contract farming arrangements, the project targets farming households with relatively greater tenure security.

[8]Two sheds that received a pilot version of the program in the previous year were excluded from the study.

2.3 Experimental design and randomization

The study design varies three major margins of a contract offered to Dunavant farmers for adoption of *Faidherbia albida*. The base parameters of the contract are a fixed number of seedlings (50, or enough to cover half a hectare) to be planted and managed by the farmer and his or her household. In keeping with the existing bonus schemes that study farmers face, and to facilitate empirical identification (see Section 4), the program offers a threshold incentive conditional on performance (tree survival) after one year.[9] Farmers were paid conditional on keeping 70 percent (35) of the seedlings alive through the first dry season (for 1 year).[10] The rationale for offering short-run incentive payments is that the costs associated with planting and caring for the trees are highest during the first year. Once young trees are established, the labor costs associated with the trees decreases substantially.

Figure 2 summarizes the experimental design. First, the size of the input subsidy (A) varied at the farmer group level in increments of 4,000 ZMK from a subsidy of 12,000 ZMK (fully subsidized) to zero. At zero subsidy, farmers paid 12,000 ZMK (approximately USD 2.60) for inputs, which is the average cost recovery price for the implementing organization. Groups were pre-assigned to one of four input subsidy treatments using the min max T approach (Bruhn and McKenzie 2009), balanced on Dunavant shed, farmer group size and day of the training. Balance results are described in Section 3.1.

Second, the size of the conditional reward (R) was varied at the individual level, in increments of 1,000 ZMK, ranging from zero to 150,000 ZMK or approximately 30 USD.[11] Variation in the conditional reward was introduced by asking participants to draw from pre-printed scratch off cards. One-fifth of all draws were for zero ZMK with the remaining four-fifths distributed uniformly over the range.

Third, the timing of the payment draw was varied at the individual level to occur either before or after the farmer's take up decision. The timing treatment was pre-determined

[9] Numerous other incentive schemes would be feasible in this setting. In addition to its advantages in our empirical design, the threshold incentive was useful for feasibility, consistency with other related contracts in development, and for ease of explanation to the farmers.

[10] The threshold was chosen based on *Faidherbia* survival rates in other programs in Eastern Zambia. Specifically, for a sample of around 3000 farmers tracked by another NGO, which offered no performance incentive, a 70 percent survival rate was achieved by around 20 percent of farmers. Thus, the threshold was set to encourage effort and minimize inframarginal payments.

[11] At the time of the study, the exchange rate was just under 5000 ZMK = 1 USD. In piloting, the distribution of payments extended to 200,000 but was scaled back prior to implementation. The scratch cards with values between 150,000 and 200,000 were removed from the prepared cards by hand, but six of them were missed. For the main analysis, we top-code payments at 150,000 and perform robustness checks that include these high outliers.

based on the registration number. Thus, in the known reward timing treatment, farmers were aware of the existence of the conditional payments and when they made their decision to take up the program, while in a surprise reward treatment farmers learned about the payments and realized their draw only after joining the program. This timing variation to separate selection effects from effort effects is similar in spirit to Karlan and Zinman (2009). We do not control for or measure beliefs about potential financial benefits from joining in the surprise reward treatment.

2.4 Data and implementation

The field experiment was implemented between November 2011 and December 2012. The study included the following stages: farmer training, program enrollment, baseline survey, end line survey, tree monitoring, reward payment. We describe each of these stages in turn.

Farmer training

Lead farmers attended a training at the Dunavant shed, which informed their own farmer group training sessions.[12] Study enumerators scheduled farmer group trainings in advance and attended the group training sessions. Group trainings were held approximately four weeks after the lead farmer training, and either the household head or an immediate family member attended, as per Dunavant's usual practices. At the group training, farmers were provided with instructions on planting and caring for the trees, were given information about the private fertilizer benefits and public environmental benefits of the trees, and were informed about the eligibility requirements for the program.[13] All farmers who attended the group training received a show up fee of ZMK 12,000 and lunch.[14]

[12]As described above, this is standard practice. Study enumerators collected basic information from the lead farmers during the training to inform the randomization. In addition to training their own farmers, lead farmers were responsible for constructing a nursery in their community, and were paid for the labor and materials. Backup nurseries were also constructed at the Dunavant sheds.

[13]Eligibility required that land must have been un-forested for 20 years, must be under the ownership of the farmer, and must not be under flood irrigation. Study enumerators confirmed that all lead farmers covered the necessary training material and summarized the main points at the end of the training to ensure information was available to all farmers.

[14]The show up fee is equal to approximately a day's agricultural wages. Note that the show up fee was sufficient to cover the cost of inputs at even the lowest subsidy level. This allows us to abstract from liquidity constraints in analyzing the effect of the subsidy on take up.

Program enrollment and baseline survey

After the lead farmer completed the group training, study enumerators explained the details of the contract, including the cost of the inputs (12,000 ZMK minus the subsidy), which was announced to the group as a whole. To implement the randomization of the rewards and to record take up decisions privately, farmers were asked to form a line. Farmers were called up one-by-one and the reward and timing treatment were implemented. The survey team leader used the registration list to determine whether the farmer was in the known or the surprise reward treatment. In the known reward treatment, farmers were given the opportunity to draw a scratch off card from a bucket. After the farmer had scratched the card to reveal its value, he was asked if he would like to join the program. The surprise treatment proceeded similarly but the take up decision was recorded before the farmer was informed that conditional payments were available. After agreeing to join the program, the farmer was told about the rewards and given the opportunity to draw a scratch card. After drawing the card and revealing its value, farmers in the unknown treatment were given the opportunity to change their minds about participation. None did.

Following the take up decision, all farmers, regardless of take up, were given a baseline survey that lasted for approximately one hour. Take up decisions were recorded for 1317 farmers, and survey data collected for 1292. Contracts were signed after the survey was complete, after which farmers paid the subsidized input cost and collected their seedlings.[15]

In addition to testing for balance in treatment assignment (discussed in Section 3.1 below), Appendix table A.1 provides some summary statistics on the study sample. Around 70 percent of participants are heads of household and around 13 percent of households are female headed. Respondents have, on average, just over 5 years of education and live in households with just over 5 members. Respondents' answers to a time preference question reveals an ordinal ranking of their discount rate, where 5 represents the lowest value placed on the future. Non-agricultural assets represent an unweighted sum of durables owned by the household. Households have around 3 hectares of land spread across just under 3 fields, which are an average of around 20 minutes away from their dwelling. Around 10 percent of households state that soil fertility is one of the major challenges that their household faces. Households have worked with the partner organization for an overage of over 4 years and over 70 percent report visits from their lead farmer in the past year. Very few are affiliated with other local organizations that promote *Faidherbia albida* (CFU and COMACO). Almost 70

[15]In a small number of cases, the lead farmer's nursery had not generated enough seedlings to support the group. In those cases, seedlings from the backup nursery were transported to the site.

percent of respondents report familiarity with the technology but only around 10 percent had adopted prior to the program. Finally, the survey asked farmers to recall risks to the trees described in the training. Out of four risks described in the training (lack of water, termites, weeding, animals), respondents remembered around 1.7.

End line survey, tree survival monitoring and conditional payments

One year after program take up, farmers were visited to record household level data and tree survival outcomes, and to deliver earned rewards. These visits were organized as follows. First a team of surveyors visited each farmer group and provided information about how the tree monitoring and payments would be implemented, then conducted an end line survey with each of the farmers, regardless of their take up decision. Of the 1292 farmers in the baseline survey, 1232 (>95 percent) are also in the end line survey. Within a week of the end line survey team's visit, a team of monitors visited the group and conducted visits to participating farmers' fields. In the field, the farmer and the monitor went tree by tree, recording whether it was sick, healthy or dead. All surviving trees counted toward the payment threshold. In case of disputes, a third party was called in from the village. The farmer signed a logsheet recording the number of agreed trees, and the monitor recorded the field's perimeter using a handheld GPS device. Additional data to proxy for quality of care, such as tree height and diameter and evidence of weeding and watering were recorded.[16]

[16]Given randomization of rewards at the individual level, there arises a potential concern about reselling seedlings to those with higher rewards or transplanting young trees just before monitoring. We use several pieces of information to investigate these concerns. First, we would expect that take up would reflect the potential to re-sell if farmers were aware of the individual-level variation in rewards. In other words, farmers that were aware of the arbitrage opportunities generated by the variation in the reward treatment should be more likely to take up, and increasingly so as the size of the subsidy falls (i.e. would respond less to the subsidy). However, a regression of take up on the interaction of the surprise reward treatment and the level of the input subsidy shows no significant interaction or clear pattern of coefficients. Second, to investigate the potential for transplanting, we take advantage of planting data collected for all farmers shortly after the end of the rainy season. We construct a measure of the difference between the planting and the monitoring tree counts, which is positive for around 100 farmers. The positive value indicates either very delayed planting or transplanting. Restricting attention to those with a positive value, the coefficient from regressing this measure of extra trees on the size of the reward is insignificant, and becomes negative (and insignificant) when group fixed effects are included. Third, we examine the within-group spillovers associated with the effect of the reward on tree survival outcomes. To the extent that transfers of any kind are happening within group, we expect a steeper slope on the reward within-group than on average. We observe a slightly smaller and statistically indistinguishable coefficient on the reward when group fixed effects are included, relative to the coefficient without fixed effects. We also test whether a farmer's own tree survival outcome depends on the reward received by others in the group. If transplanting or resale affect a large share of outcomes, we expect to see own performance falling in the relative size of others' rewards. We see the opposite: to the extent that any reward spillovers are present, they are positive, with tree survival increasing in the size of

Finally, within a couple of days of the monitoring visit, the reward payment team visited each group and paid out rewards to farmers with 35 or more surviving trees. Keeping the payments separate from the monitoring was intended to improve monitors' objectivity.[17] Of the farmers eligible for monitoring, 9 could not be found for monitoring. We assume zero tree survival for these individuals in our main analyses and will include robustness checks on the sensitivity of our results this assumption in future drafts.

3 Summary statistics and reduced form results

Before moving to the structural model, we summarize the data and present reduced form results. We begin by empirically testing the validity of the experimental design, then present the effect of the experimental treatments on contract take up and performance. We focus, in the reduced form analysis, on evidence for selection on type and discuss the implications for unobserved heterogeneity, which we explicitly estimate in Section 4.

3.1 Balance and attrition

Reduced form identification of treatment effects requires that treatment assignment is balanced along observable (and therefore unobservable) characteristics. We show balance statistics for both the input subsidy and reward levels, as well as the surprise reward treatment in Appendix table A.1. Columns 1, 3 and 5 report the means and standard deviations for the zero input subsidy, zero reward and known reward timing treatment, respectively. Column 2 shows the coefficients and standard errors from a linear regression of the covariate on input subsidy treatment dummies. Column 4 reports coefficients and standard errors from univariate regressions of each of the covariates on the reward level, and column 6 reports the same with the surprise reward treatment as the righthand side variable. We observe some imbalance in the assignment of the input subsidy treatment with slightly larger households with more non-agricultural assets receiving lower input subsidy assignments on average. We also observe that older respondents with larger households and better self-reported soil fertility

others' rewards. This result is consistent with effort spillovers but not with resale or transplanting.

[17] As a check for collusion between the monitors and farmers, we test whether individual monitors are associated with a higher probability that a farmer passes the tree survival threshold. No single monitor indicator is significantly correlated with reaching the threshold, nor are the monitor indicators jointly predictive. This implies that either all monitors were cheating to the same degree or that no monitors were cheating. Given differences in career concerns with the study implementers (some had higher paid jobs as survey supervisors when not engaged in monitoring), we believe that incentives to cheat varied across farmers.

are marginally more likely to be assigned to the surprise reward treatment. The table tests balance for 17 variables and consists of 51 separate regressions. Five coefficients that are significant at $p < 0.10$ is therefore expected. Our results include these variables as controls.

Our identification strategy will be violated if selection into the take up decision is affected by treatment assignment, if attrition is correlated with treatment or if spillovers across treatment groups affect take up or implementation. To address the first concern, treatments assigned at the group level were not revealed until after individuals arrived for training, making selection into take up on the input subsidy treatment unlikely. Conditional on attending the training, only one take up observation is missing.

To address the second concern, we examine attrition at each stage of data collection in Appendix table A.2. Column 1 reports means and standard deviations of the treatment variables for all 1317 individuals who made a take up decision. Columns 2, 3 and 4 show coefficients and standard errors for univariate linear regressions of each variable regressed on a dummy for participation in the baseline survey, end line survey and in the tree monitoring, respectively. Participation rates in the baseline are high, with over 98 percent of individuals included. For the end line, 94 percent of all study farmers and over 95 percent of baseline respondents were located. We see some evidence that farmers who received lower input subsidies were marginally less likely ($p < 0.10$) to participate in the surveys. Otherwise, attrition is balanced across treatments. For the tree monitoring, over 95 percent of the 1092 households that took up the program were located. The tree monitoring is the source of our primary outcome measure and, reassuringly, we see no evidence of non-random attrition at the monitoring stage.

Finally, the third concern is related to spillovers across treatments. Because the input cost treatment was assigned at the group level, spillovers are relatively unlikely. Indeed, during regular effort monitoring visits, no farmers asked about difference in input costs. The value of the reward, on the other hand, varied at the individual level within-group. By revealing the reward treatment one by one immediately before soliciting the take up decision, we mitigate the potential that take up is affected by rewards received by others.[18] We observe some suggestive evidence that being in a farmer group with an average higher reward draw has a positive spillover on own survival probabilities, though the magnitude of the effect is small relative to the direct effect of the reward.

[18]This is confirmed empirically by regressing the probability of take up on the average random reward draw that preceded a farmer's own draw. Earlier draws have no effect on the probability of take up.

3.2 Participation

We measure the participation response to the exogenous variation in the contract parameters by estimating the linear regression

$$y_{ig} = \alpha + \beta A_g + \delta R_i + \phi X_i + u_{ig} \tag{1}$$

where y_{ig} is the binary participation decision for farmer i in group g, A_g is the input subsidy applied to group g, R_i is the reward offered to farmer i conditional on keeping 70 percent of his trees alive for one year and X_i is a vector of individual level controls. Errors are clustered at the farmer group level g. All specifications also control for the random sample of farmers who receive additional data collection visits (effort monitoring) during the course of the year and the timing of the reward announcement.

We estimate equation 1, with y_{ig} equal to 1 if the farmer takes up the program and zero otherwise, and include an interaction term for A_g and R_i. Table 1 shows the results of linear regressions of the binary take up outcome on the input subsidy and reward conditions, and their interaction (equation 1). Participation is increasing in the input subsidy (column 1), with take up rates close to 100 percent when inputs are fully subsidized. Relative to farmers offered the contract at a full subsidy, subsidies of 8,000, 4,000 and zero ZMK decrease participation by 11, 22 and 26 percent respectively. Take up is also (weakly) increasing in the linear reward (column 2), with a one thousand Kwacha increase leading to 0.06 percent higher take up. Finally, the effect of the reward differs across input subsidy conditions (column 4) with larger effects at lower subsidies. While participation is responsive to the exogenous variation in contract payoffs, the response explains relatively little of the total variation in take up. Overall, 10 percent or less of the variation in take up observed in the data is explained by the variation in subsidies and rewards, which suggests that heterogeneity across farmers plays an important role in the decision to participate.

3.3 Incentive effect of reward payments

The distribution of tree survival outcomes is informative of both the effect of the reward payments and the variance in private benefits from the trees. Given the threshold structure of the contract, tree production outcomes other than zero or the performance threshold are driven by private payoffs from the trees. Figure 3 shows the histogram of tree survival outcomes for zero and positive reward levels. Farmers who did not take up are included with a tree survival outcome of zero. The distribution of outcomes when the reward is zero

provides a counterfactual description of tree survival outcomes in the absence of reward payments. Relative to tree production in the absence of a reward, the number of farmers producing zero trees is substantially smaller when the reward is positive. Positive rewards shift mass to the right of the threshold.

Table 2 shows linear regression estimates at several tree survival outcome intervals. Panel A includes only farmers who did not select on the reward, namely those in the surprise reward treatment.[19] Panel B includes all farmers, with tree survival outcomes set equal to zero for farmers who did not take up the program. In Panel B we interact a dummy for a non-missing reward value with the reward in '000 ZMK to avoid dropping farmers in the surprise reward treatment who did not take up the program.

Columns 1 and 2 of Panel A show that the average number of surviving trees increases by 0.07 (s.e. 0.02) for each thousand Kwacha increase in the reward, or by about 7.3 (s.e. 1.8) trees when rewards go from zero to the average positive reward of around 77,000 ZMK. The probability of having zero trees is decreasing in the reward by 0.1 percent (s.e. 0.0004) for every thousand Kwacha increase in the reward (column 3). The probability of keeping at least the threshold number of trees alive is increasing in the reward (columns 5 and 6), yet the number of trees produced, conditional on being above the threshold is not responsive to the reward (columns 9 and 10), consistent with production above the threshold being driven by private benefits. The results are similar, though slightly smaller in magnitude, for the full sample shown in Panel B, which includes the selection effect of the reward for farmers who made their participation decision after drawing R.

Together, these tree survival results suggest a large amount of variance in the private payoffs associated with tree production, which can be seen most clearly in the top panel of Figure 3. At the same time, conditional rewards have a significant effect on survival outcomes, so this variance is not associated with a lack of control over tree survival outcomes.

3.4 Selection effects

Our design separately identifies the selection induced by the subsidy on inputs and the reward on outcomes. In our context, selection is defined as the differences in participant expected private payoff from the contract across treatments. Specifically, as the level of subsidy or reward for outcomes decreases, the expected private payoff needed to induce participation increases. With little or no uncertainty about payoffs, higher expected private

[19] We allow for more flexibility in the effect of rewards on tree survival in the next sub-section.

payoff individuals will keep more trees alive, i.e. we will observe a selection effect in farmers' performance.

Thus, as the input subsidy (A) increases, the average private benefit of participants falls, resulting in lower average tree survival due to this selection effect.[20] Whether or not we observe an effect of selection in response to the input subsidy on tree survival outcomes depends on farmers knowing their private payoffs. If the take up decision is made under sufficient uncertainty, then the participation decision, which reflects the input subsidy, may be a poor predictor of tree survival outcomes. Conditional on take up, Equation (1) estimates the selection effect of A_g on tree survival, and the combined selection and incentive effect of R_i on tree survival. We assume that all non-takers have tree survival outcomes of zero and estimate the intention to treat effects, which allows for a causal interpretation of the combined selection and incentive effects of the input subsidy. Figure 4 plots average tree survival from regressions of tree survival on the input subsidy treatments, controlling for the rewards, monitoring condition and surprise reward treatments. Conditional on take up, tree survival is constant across all four input subsidy conditions, suggesting little effect of A on self-selection. For the intention to treat results, the participation effect of A is combined with the selection effect, and given the lack of selection effect, results are similar to those shown in Table 1.

Analysis of the effect of selection on the reward on tree survival is less straightforward, since R enters into both the farmer's participation decision and tree survival outcome. To isolate the causal effect of the reward on tree survival separately from its effect on selection in to the program, we interact a treatment indicator for the surprise reward treatment with the level of the reward:

$$trees_{ig} = \alpha + \beta A_g + \delta R_i + \gamma surprise_i + \lambda R_i \times surprise_{ig} + \phi X_i + u_{ig} \tag{2}$$

where $\delta + \lambda$ identifies the causal effect of the reward on tree survival, conditional on take up. Specifically, recall that farmers in the surprise reward treatment decided to join the program before learning about the reward, so the effect of R on tree survival for these farmers comes solely from the incentive effect (see also Table 2). For farmers in the known reward treatment, who drew R before they decided whether or not to participate, the effect of R on tree survival

[20]This assumes that the amount paid for inputs does not directly enter into the farmer's optimal number of trees, which is not valid in the presence of behavioral biases. The sunk cost fallacy has been studied in multiple developing country field experiments (Ashraf et al. 2010; Cohen and Dupas 2010; Berry et al. 2012), none of which has found evidence of its empirical relevance. In our setting, a sunk cost effect and a selection effect work in the same direction and so can both be ruled out in the case of a null result.

is made up of both the selection and incentive effects. Therefore, the difference between the known and the surprise reward treatment (λ) isolates the selection effect of R, conditional on take up. We would expect λ to be negative if higher rewards attract lower net benefit types.

Figure 5 shows marginal effects of tree survival on the reward timing treatment interacted flexibly with R_i in four intervals: $\{0, 1 \text{ to } 50, 51 \text{ to } 100, 101 \text{ to } 150\}$, conditional on take up. The lines are overlapping and statistically indistinguishable at all reward levels. The figure therefore clearly demonstrates that a higher reward does not induce self-selection that translates into meaningful tree survival differences.

Together, the two selection tests indicate that the participation decision is a poor predictor of tree survival. This finding is consistent with a large share of the variation in payoffs from the contract that is determined by variation in the realized shocks to effort cost. The high average participation is also consistent with the importance of the option value conveyed by the contract in the absence of penalties for non-compliance.

4 Model, identification and estimation

We turn now to a model of intertemporal decision making under uncertainty, in which agents respond to a threshold contract for tree production. The reduced form results showed little evidence for selection on agent type in our setting, which suggests that agents know relatively little about their expected profits under the contract at the time of take up. The structure imposed by the model allows us to explicitly estimate the distribution of both the contract payoffs that are known to the agent ex ante and net benefits that are realized ex post. The magnitude of the uncertainty faced by the agent at the time of the participation decision determines the amount of selection we observe as well as the role that option value – the ability to exit the contract penalty-free at any point after take up – plays in driving take up. We describe the model and how it is identified from the exogenous variation in contract parameters, and the estimation procedure.

4.1 Farmer net benefits

General profit function We start by considering the decision of a farmer to plant and care for trees according to the private benefits they offer. Consider the general quadratic profit function given by

$$\Pi(N) = \sum_{t=5}^{\bar{T}} \left[\frac{1}{(1+r)^t} \left(\alpha_0 N - \alpha_1 N^2 \right) \right] - \gamma_0 N - \gamma_1 N^2 - \gamma_2 \times \mathbf{1}(N > 0) \tag{3}$$

where α_0 and α_1 are private benefit parameters, r is the annual discount factor, $\bar{T}$ is the maximum number of years a tree lives, γ_0 and γ_1 are cost of implementation parameters that are proportional to the number of trees and γ_2 is a fixed cost of implementation.

We can rewrite (3) as

$$\Pi(N) = (\delta\alpha_0 - \gamma_0) N - (\delta\alpha_1 + \gamma_1) N_1^2 - \gamma_2 \times \mathbf{1}(N > 0)$$

where $\delta = \sum_{t=5}^{\bar{T}} \left[\frac{1}{(1+r)^t} \right]$. This new expression for the profit makes clear that, with $\alpha_0 > 0$, $\alpha_1 > 0$, $\gamma_0 > 0$, $\gamma_1 > 0$, $\gamma_2 > 0$, and $\delta\alpha_0 - \gamma_0 \geq 0$, equation (3) essentially assumes decreasing marginal value to the number of trees planted and cared for. This is consistent with our empirical observation that a number of farmers plant a strictly positive number of trees that is above zero and below 50, the number of seedlings they receive, whenever the reward for tree survival is equal to zero.

Once all implementation costs and benefits are revealed, a farmer chooses the number of trees to plant and cared for that maximizes (3). The optimal number of trees planted and cared for by farmers is characterized by

$$N^* = \begin{cases} \frac{\delta\alpha_0 - \gamma_0}{2(\delta\alpha_1 + \gamma_1)} & \text{if } \frac{(\delta\alpha_0 - \gamma_0)^2}{4(\delta\alpha_1 + \gamma_1)} - \gamma_2 > 0 \\ 0 & \text{if } \frac{(\delta\alpha_0 - \gamma_0)^2}{4(\delta\alpha_1 + \gamma_1)} - \gamma_2 \leq 0 \end{cases}$$

Note that in principle, all parameters in (3), $\alpha_0, \alpha_1, \delta, \gamma_0, \gamma_1$, and γ_2, may vary across farmers. Our experimental variation, however, would not allow us to separately identify all potential sources of heterogeneity across farmers. We therefore turn to a quasi-profit function that allows us to use our experimental variation to characterize farmer heterogeneity along two important dimensions of the farmer's profit maximization problem, $\max_N \Pi(N)$: the interior solution and the profit level evaluated at the optimal number of trees for the farmer.

Quasi-profit function We can write a quasi-profit function indexed by two parameters that delivers the same interior and corner solution conditions as (3).

$$\Pi_Q(N|T_i, F_i) = N - \frac{(\delta\alpha_1 + \gamma_1)}{(\delta\alpha_0 - \gamma_0)} N^2 - \left[\gamma_2 - \left(1 - \frac{1}{(\delta\alpha_0 - \gamma_0)} \frac{(\delta\alpha_0 - \gamma_0)^2}{4(\delta\alpha_1 + \gamma_1)} \right) \right] \times \mathbf{1}(N > 0) \tag{4}$$

This alternative function allows for heterogeneity across farmers in the interior solution, $\frac{\delta\alpha_0-\gamma_0}{2(\delta\alpha_1+\gamma_1)}$, as well as the scale of the profit associated with the interior solution, $\left[\gamma_2 - \left(1 - \frac{1}{(\delta\alpha_0-\gamma_0)}\frac{(\delta\alpha_0-\gamma_0)^2}{4(\delta\alpha_1+\gamma_1)}\right)\right]$ For simplicity, we can reparameterize (4) as

$$\Pi_Q(N|T_i, F_i) = N - \frac{1}{2T_i}N^2 - F_i \times \mathbf{1}(N > 0) \tag{5}$$

where $T_i > 0$ is the interior solution to the optimization problem and $F_i \in (-\infty, \infty)$ is the scale plus fixed cost parameter that determines the corner solution. The solution to the optimization problem defined by $\max_N \Pi_Q(N)$ is given by

$$N^* = \begin{cases} T_i & \text{if } \frac{1}{2}T_i - F_i > 0 \\ 0 & \text{if } \frac{1}{2}T_i - F_i \leq 0 \end{cases}$$

There are two important advantages of (5) vs. (3). The first is that all random parameters are identified out of the variation induced by our experiment. The second is that the profit function and the quasi-profit function take the same value when evaluated at the optimal solution. I.e.,

$$\Pi(N^*) = \Pi_Q(N^*|T_i, F_i) = \begin{cases} \frac{1}{2}T_i - F_i & \text{for } N^* > 0 \\ 0 & \text{for } N^* = 0 \end{cases}$$

Hence, (5) can be used to evaluate welfare under the more general profit function (3).

From here on, we assume that in the absence of an external incentive or reward for tree survival, farmers maximize

$$\Pi_Q(N|T_i, F_i) = \left(N - \frac{1}{2T_i}N^2\right) - \mathbf{1}(N > 0) \times F_i$$

with respect to the number of trees, N.

Introducing a reward for survival

Farmer's objective function when they are faced with a reward for tree survival is assumed to be

$$\Pi_R(N|F_i, T_i, R_i) = \left(N - \frac{1}{2T_i}N^2\right) + \mathbf{1}(N \geq \bar{N})R_i - \mathbf{1}(N > 0)F_i \tag{6}$$

where R_i denotes the reward for reaching the goal of $\bar{N}$ trees. A farmer that faces a reward of R_i will maximize her objective function (6) subject to the choice of trees N. Denote N_i^{**} the number of trees that maximizes (6),

$$N_i^{**}(T_i, F_i, R_i) = \arg\max_N \left(N - \frac{1}{2T_i} N^2 \right) + \mathbf{1}(N \geq \bar{N}) R_i - \mathbf{1}(N > 0) F_i \tag{7}$$

4.2 Dynamics, timing of information, and program participation decision

In order to analyze the effects of uncertainty about net benefits on the performance of the incentive program, we assume the farmer makes program-related decisions in two periods of time. In an early period, denoted $t = 0$, the farmer decides whether or not to pay the cost of the tree seedlings in order to participate in the tree planting program. At this point in time, the farmer has partial information about her net benefits from the contract. At a later point in time, denoted $t = 1$, the farmer makes the decision of how many trees to plant and care for described by (7). At this point in time, the farmer-specific determinants of their net benefit function are fully known to the farmer.

More specifically, we separate the scale of the net benefit into two components such that $F_i = F_{0i} + F_{1i}$. Therefore, we now have three farmer-specific determinants of the profit function: T_i, F_{0i}, and F_{1i}. We assume that out of these three determinants, only F_{0i} is known to the farmer at the time the participation decisions is made ($t = 0$), while F_{1i} and T_i are revealed to the farmer at $t = 1$. Finally, we assume that the farmer knows the the distribution of F_{1i} and T_i conditional on F_{0i} at the time she makes her decision to participate.

Given this timing for the information, a farmer decides to participate in the program if

$$\mathbb{E}\left[\max_N \Pi_1(N | T_i, F_{0i}, F_{1i}, T_i, R_i) \middle| F_{0i} \right] - c + A_i > 0 \tag{8}$$

where c is the cost of the seedlings and A_i is the randomly determined subsidy. The expectation in (8) is taken over the distribution of (F_{1i}, T_i) conditional on F_{0i}.

4.3 Identification and estimation

Identification of the model depends on our ability to estimate the distributions of farmer-specific net benefit components, F_{0i}, F_{1i}, and T_i, which are unobserved to the researcher. Identification relies on

a) the observed variation in N_i^{**}

b) random variation in R_i

c) random variation in A_i

d) a distributional assumption for F_i and T_i conditional on F_{0i}.

1. The distribution of T_i, G_T, is identified from the observed choices of N_i^{**} when $R_i = 0$, and when $N_i^{**} > \bar{N}$ and $R_i > 0$. Note that $N_i^{**} = T_i$ when $R_i = 0$ and $\frac{1}{2}T_i > F_i$.

2. The distribution of $F_i = F_{0i} + F_{1i}$ is identified out of the shifts in probability mass from $\Pr(N^{**} = 0)$ to $\Pr(N^{**} \geq \bar{N})$ and from $\Pr\left(0 < N^{**} < \bar{N}\right)$ to $\Pr(N^{**} = \bar{N})$, as R increases.

3. The variance of F_{0i}, $\sigma^2_{F_0}$, is identified from the relationship between the decision to participate and A_i and the following assumptions

   (i) $F_{0i}, F_{1i} \perp T_i$

   (ii) $F_{0i} \sim n\left(\mu_F, \sigma^2_{F_0}\right)$, $F_{1i} \sim n\left(0, \sigma^2_{F_1}\right)$, $F_{1i} \perp F_{0i}$, and

   (iii) the farmers know F_{0i} but they know neither F_{1i} nor T_i at the time they make their decision to participate.

To see this note that the decision to participate is determined by the inequality

$$\mathbb{E}\left[\max_N \Pi_1(N|T_i, F_{0i}, F_{1i}, T_i, R_i)\middle| F_{0i}, R_i\right] > c - A_i \tag{9}$$

The left side of the above inequality is a known function of the random variable F_{0i} up to one parameter, $\sigma^2_{F_0}$. Denote this function $h(F_{0i}; \sigma^2_{F_0}, R_i)$. The right side of the inequality is known to take one of four values, $c - a_j$ for $j = 1, ..., 4$, and is randomly determined by the research design. Provide that $h(F_{0i}; \sigma^2_{F_0}, R_i)$ is invertible, we can identify parameter, $\sigma^2_{F_0}$ which governs the distribution of F_{0i}, from the four moments defined by

$$\Pr\left(F_{0i} < h^{-1}(c - a_j; \sigma^2_{F_0}, R_i)|\mu_F, \sigma^2_F, G_T, \sigma^2_{F_0}, R_i\right) = \Pr(Part_i|A_i = a_j) \text{ for } j = 1, ..., 4 \tag{10}$$

where $\sigma^2_F = \sigma^2_{F_0} + \sigma^2_{F_1}$. Note that in μ_F, σ^2_F, G_T can be treated as known in (10) since they are identified from tree survival, as described in 1 and 2, and R_i and A_i are known by the researcher.

Estimation

We estimate two versions of the model described above. The first version assumes that all farmers face the same distribution of F_{i0}, F_{i1} , and T_i. The second version allows these distributions to vary with observable characteristics of the farmer.

For the estimation of the first version of the model, we further adopt a parametric assumption on T_i. Because the number of trees that survived look somewhat evenly distributed across outcomes, we assume $T_i \sim \text{uniform}(0, b]$.[21]. We also assume that F_{1i} and F_{0i} are both normally and independently distributed with $F_{1i} \sim n(0, \sigma_{F_1})$ and $F_{0i} \sim n(\mu_F, \sigma^2_{F_0})$. We also assume that $F_{0i}, F_{1i} \perp T_i$.

We estimate the model using simulated maximum likelihood. The log-likelihood function will be over observations of the number of planted trees, $N = 0, ..., 50$, and the participation decision, $P = 0, 1$. Because there are no trees planted whenever the individual chooses not to participate, the support of this bivariate vector is given by the 51 (P, N) pairs: $(0, 0), (1, 0), (1, 1), (1, 2), ..., (1, 50)$.

$$l(\xi; DP, \tilde{N}) = \sum_{i=1}^{M} \left\{ (1-DP_i) \ln(1-\pi_{P,i}) + DP_i \ln(\pi_{P,i}) + DP_i \sum_{j=0}^{50} I(N_i = j) \ln \Pr(N = j) \right\}$$

where $\xi = (\mu_F, \sigma^2_{F_0}, \sigma^2_{F_1}, b)$.

In the second version of the model, we assume that $F_{0i} \sim n(\mu_{Fi}, \sigma_{F_0})$, $F_{1i} \sim n(0, \sigma_{F_1})$, and $T_i \sim \text{uniform}(\exp(X\beta_T))$, where $\mu_{Fi} = X_i\beta_F$ and $b_i = \exp(X_i\beta_T)$. This allows $F_i = F_{0i} + F_{1i}$ and T_i to be correlated through observable characteristics of the farmer. Hence, in the second version of the model the parameter set becomes $\xi = (\beta_F, \sigma^2_{F_0}, \sigma^2_{F_1}, \beta_T)$.

We use numerical methods to minimize the negative of the log-likelihood. For each likelihood evaluation, we use 1000 draws of (T_i, F_{0i}, F_{1i}). Also within each likelihood evaluation and for each draw of (T_i, F_{0i}, F_{1i}), the expectation in the right hand side of (9) is numerically computed conditional on the draw of F_{0i}.[22]

[21] We will experiment with other distributions such as truncated normal.

[22] Simulated methods often result in stepwise objective functions which work poorly with gradient-based numerical optimization algorithms. To facilitate the numerical optimization, we "smooth" the objective function by computing the multilogit formula for each decision over participation and the number of trees. We assume a relatively small variance parameter of the logistic error term: 0.5. However, we experiment with different values for this parameter. We find that smoothing does not significantly affect the point estimates and does improve substantially the curvature of our objective function.

5 Welfare

In this section, we describe the approach to investigating the welfare implications of our model. Results are pending.

5.1 Welfare

Next, we consider the social benefits of the program by accounting for private and public benefits in a single welfare function. Denote α as the social value of carbon, η as the marginal cost of public funds, and m as the monitoring cost. The average welfare per farmer conditional on reward R and subsidy A is thus

$$W(A,R) = \mathbb{E}_{F_{0i}}\left[\Pr(Part_i|F_{0i};A,R)\left(B(F_{0i};A,R) - C(F_{0i};A,R)\right)\right] \tag{11}$$

where benefits are equal to

$$B(F_{0i};A,R) = \alpha \times \mathbb{E}_{T_i,F_{1i}|F_{0i}}\left[N^{**}(T_i,F_{0i},F_{1i};R)\right] + \mathbb{E}_{T_i,F_{1i}|F_{0i}}\left[\Pi_R\left(N^{**}\left(T_i,F_{0i},F_{1i};R\right)\right)\right],$$

costs are equal to

$$C(F_{0i};A,R) = R \times \Pr\left(N^{**}\left(T_i,F_{0i},F_{1i};R\right) \geq \bar{N}\middle|\, F_{0i};R\right) \times (\eta) + [A + m \times \mathbf{1}(R>0)] \times (\eta - 1)$$

and f(F_{0i}) is the marginal distribution of F_{0i}.

Equation 11 offers some intuition for the tradeoffs associated with R and A. On the one hand, the social planner incurs public funds costs associated with R (ηR) only if individuals comply (i.e. $N^{**} > \bar{N}$), while the costs associated with the subsidy ($(\eta-1)A$) are incurred for everyone that participates. This makes R an attractive policy instrument as compared to A in the presence of uncertainty, provided that monitoring costs are small or nonexistent. In the presence of monitoring costs, m, R becomes less attractive as a policy instrument. This is because monitoring costs are incurred for everyone that participates, and the presence of R will increase participation rates of all individuals, including those with a very low probability of complying.

Also, note that together with participation rates, R increases the expected net private benefit of the contract, $\mathbb{E}\left[\Pi_R\left(N^{**}|F_{0i},F_{1i},T_i,R\right)|F_{0i},R\right]$. Moreover, the effect is higher, the higher the uncertainty adopters face. Hence, R further increases the private component of welfare in the presence of uncertainty.

Using our model estimates, we characterize the optimal combination of A and R in our setting. We then examine how sensitive is this combination to changes in monitoring costs, as well as different levels of uncertainty.

6 Alternative explanations

We explore three potential explanations that match our empirical results, but are not explicitly reflected in our model: common shocks, procrastination and learning.

6.1 Common and idiosyncratic shocks

By interpreting the ex post variance in profits across farmers as the sum of the variances of the known and unknown components of the profits, our model allows for idiosyncratic but not common shocks. In all likelihood, farmers face both common and idiosyncratic shocks and to the extent that the former is important, it will generate downwards bias in the variance parameter that describes the unobserved shocks.[23] That said, heterogeneity in the exposure to common shocks will reduce the correlation of ex post profits across farmers and will therefore allow us to partially capture common shocks through the unknown component of the profits, i.e. the idiosyncratic shock term. We consider two sources of information about the relative importance of common and idiosyncratic shocks: household self-reports from the baseline and endline data and the existing literature on agricultural productivity in rural Sub-Saharan Africa.

First, the two most often discussed common shocks are weather and prices. For example, the primary source of price risk is the price on cotton, which experienced a very negative shock in the year of study. At baseline, 97 percent of respondents who forecasted a minimum possible price on cotton for the coming year exceeded the realized price. Close to 80 percent of the households in our study grew cotton in the contract year, and were therefore affected by the price shock. However, the share of a household's agricultural land under cotton ranges from zero to 100 percent, with a mean of 34 percent and a standard deviation of 26 percent. Exposure to cotton price risk therefore varies substantially across households.

[23]The assumption that ex post variance in profits is the sum of the variances of the known and unknown components of the profits is equivalent to assuming that shocks are uncorrelated across farmers. A positive correlation in the shocks across farmers will result in a lower ex post variance. Hence, the model will assume that the residual variance (the left over variance once the variance in known shocks is subtracted) corresponds to the variance of the unknown shocks. This restriction will cause an underestimation of the variance in the unknown component of the shocks.

Weather represents the other likely source of common risk exposure. Unlike cotton prices, rainfall was fairly typical during the year of the contract, according to local agronomists.

When asked about the greatest challenges facing their households, two-thirds of respondents indicated health problems, which are likely to arrive as idiosyncratic shocks, while only 2 percent indicated weather or rainfall. Further evidence of unexpected shocks are reported at endline, with almost 50 percent of households reporting lost cattle or livestock to death or theft during the past year and 10 percent of households reporting loss of a working age member to death or marriage during the year.

Second, a growing literature documents large negative effects of household illness on agricultural productivity in the region of study (Fink and Masiye 2012; Baranov et al. 2013) and on consumption more broadly (for example, Gertler and Gruber 2002). These findings are consistent with a larger literature that documents, in most cases, a disproportionate share of income risk from idiosyncratic factors in rural developing country settings (summarized in Dercon 2002). Thus, while common shocks are potentially an important source of uncertainty in agriculture, restricting our attention to idiosyncratic shocks appears reasonable in this context.

6.2 Procrastination

An alternative explanation for the observed pattern of high take up and low compliance is procrastination or hyperbolic time preferences. Sustained effort choices are frequently associated with time inconsistent behavior, in which the individuals initially adopts, intending to follow through. But when the time comes to act, costs loom larger (or benefits smaller) than anticipated at the time of initial adoption. Mahajan and Tarozzi (2011) document time inconsistent technology adoption for insecticide treated bednets in India, with low rates of net re-treatment.

We examine evidence for procrastination or hyperbolic time preferences, relying on intuitive patterns in the data rather than precise tests that benefit from exogenous variation. We construct two measures of procrastination from the survey data, the first of which relies on a series of questions asked at baseline about the respondent's tendency to spend money quickly or delay purchases or actions. These questions are combined into a binary measure of procrastination, which is likely to miss naive procrastinators.[24] As part of the endline survey,

[24]Specifically, the baseline questions prompted (1) If I get money, I tend to spend it too quickly, (2) I often change my mind and do not follow through with my original intention and (3) I tend to postpone activities until later. All responses were on a four point Likert item from strongly disagree to strongly agree

a series of questions about procrastination in other activities, including paying school fees, purchasing agricultural inputs and milling maize were added to the survey. These are combined into a binary measure if the respondent reports procrastination on any of these other activities, which is more likely than the other measure to capture naive procrastinators.

We begin by examining whether these measures of procrastination are correlated with contract take up or tree survival conditional on take up, controlling for other characteristics. They are not. We next investigate the insight that farmers prone to procrastination may be differentially sensitive to a contract structure that requires them to pay more upfront for inputs if the potential rewards arrive only after a year of effort. We regress take up on an interaction of each of the two procrastination measures and the input subsidy level. For the self-described procrastinators, there is a weakly greater likelihood of take up at higher subsidy levels. However, the interaction is insignificant with the measure more likely to capture naive procrastinators. These results are summarized in Appendix table A.3, and suggest a relatively minor role for procrastination in driving take up or contract compliance outcomes.

6.3 Learning

We interpret the model as one of intertemporal shocks to opportunity cost. If, instead, farmers are simply unaware of the costs and benefits of the trees ex ante, then learn about the technology during the first year of implementation, we may misattribute this learning to the arrival of an idiosyncratic shock. Our study design cannot clearly distinguish the two explanations, and while we provide evidence for the prevalence of shocks above, we can also explore the extent to which information changes among adopting farmers over the course of the year-long contract.

Our most reliable measure of knowledge about technology that is likely to affect tree survival outcomes is the number of risks to tree survival that the farmer is able to list. Out of four major risks to the trees (see Section 2.4), farmers list an average of 1.64 risks at baseline and 1.75 at endline one year later. There is, however, a positive effect of the difference in reported risks between baseline and endline among farmers who took up the contract. This is driven both by actual knowledge acquisition among those who gained experience with the technology through the contract, and less forgetting about risks relative to those who did not adopt. Conditional on take up, tree survival is also higher among those who had planted

and summed to create a Likert scale from zero to 12. A binary measure was constructed for responses above the scale median.

Faidherbia at baseline and presumably had less to learn about the technology. These correlations should not be interpreted as causal, but do suggest that some of the information that arrives after take up may be through learning about the technology. In future extensions to the structural model, we will simulate results under alternative assumptions about learning.

7 Conclusion

We present the findings from a field experiment that investigates the determinants of adoption of a technology that generates positive externalities. In Zambia, farmers decide whether to adopt a nitrogen fixing tree under considerable uncertainty about the costs of implementation. The study introduces exogenous variation into the initial subsidy for inputs and the size of a reward conditioned on tree survival. We observe a weak correspondence between the initial decision to take up and the the decision to follow through, which does not systematically vary across subsidy conditions. At the same time, farmers are responsive to the size of the reward in their decision to keep trees alive. Results are consistent with the arrival of shocks to the cost of implementation that arrive after the farmer has decide whether to participate in the project.

The experimental variation is used to identify a structural model of intertemporal decision making under uncertainty, which explains our field results and quantifies the uncertainty that the farmers face at the time of take up. The model offers additional intuition for the tradeoffs posed by input subsidies and outcome rewards under uncertainty. In future work, we will use the model to estimate the distribution of farmer types and ex post shocks to implementation costs, and to simulate welfare under alternative assumptions about information and costs.

References

Akinnifesi, Festus K, OC Ajayi, G Sileshi, Paxie W Chirwa, and Jonas Chianu, "Fertiliser trees for sustainable food security in the maize-based production systems of East and Southern Africa: A review," *Agronomy for sustainable development*, 2010, *30* (3), 615–629.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias, "Targeting the Poor: Evidence from a Field Experiment in Indonesia," *American Economic Review*, 2012, *102* (4), 1206–1240.

Alix-Garcia, J, K Sims, Patricia Yanez-Pagans, V Radeloff, and E Shapiro, "Two Dimensional Evaluation: The Environmental and Socioeconomic Impacts of Mexico's Payments for Hydrological Services Program," *Unpublished manuscript*, 2012.

Arrow, Kenneth J and Anthony C Fisher, "Environmental preservation, uncertainty, and irreversibility," *The Quarterly Journal of Economics*, 1974, *88* (2), 312–319.

Ashraf, Nava, James Berry, and Jesse M. Shapiro, "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia," *American Economic Review*, 2010, *100* (5), 2383–2413.

Baranov, Victoria, Daniel Bennett, and Hans-Peter Kohler, "The Indirect Impact of Antiretroviral Therapy," *Mimeo*, 2013.

Barnes, R.D. and C.W. Fagg, "Faidherbia albida: Monograph and annotated bibliography," Technical Report, Oxfor Forestry Institute, University of Oxford 2003.

Berry, James, Greg Fischer, and Raymond P. Guiteras, "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana," *Unpublished manuscript*, 2012.

Besley, Timothy and Stephen Coate, "Workfare versus Welfare: Incentive arguments for work requirements in poverty-alleviation programs," *American Economic Review*, 1992, *82* (1), 249–261.

Brown, S., "Estimating biomass and biomass change of tropical forests: A primer," Technical Report Paper no. 134, Food and Agriculture Organization of the United Nations, Rome, Italy 1997.

Bruhn, Miriam and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 200–232.

Cacho, Oscar J, Robyn L Hean, and Russell M Wise, "Carbon-accounting methods

and reforestation incentives," *Australian Journal of Agricultural and Resource Economics*, 2003, *47* (2), 153–179.

Chassang, Sylvain, Gerard Padro i Miquel, and Erik Snowberg, "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments," *American Economic Review*, 2012, *102* (4), 1279–1309.

Cohen, Jessica and Pascaline Dupas, "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, 2010, *125* (1), 1–45.

Conning, Jonathan and Christopher Udry, "Rural Financial Markets in Developing Countries," in R. Evenson and P. Pingali, eds., *Handbook of Agricultural Economics*, Vol. Volume 3, Elsevier, 2007, pp. 2857–2908.

Conte, M. N. and M. Kotchen, "Explaining the Price of Voluntary Carbon Offsets," *Climate Change Economics*, 2010, *1* (2), 93–111.

Dercon, Stefan, "Income risk, coping strategies, and safety nets," *The World Bank Research Observer*, 2002, *17* (2), 141–166.

Dixit, Avinash K. and Robert S. Pindyck, *Investment under uncertainty*, Princeton University Press, 1994.

Dupas, P., "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment," *Econometrica*, 2013, *forthcoming.*

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer, and Alix Peterson Zwane, "Free provision, micro ordeals, and targeting for preventive health," *Working Paper*, 2013.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf, "Optimal mandates and the welfare cost of asymmetric information: Evidence from the UK annuity market," *Econometrica*, 2010, *78* (3), 1031–1092.

_ , _ , Stephen P Ryan, Paul Schrimpf, and Mark R Cullen, "Selection on Moral Hazard in Health Insurance," *The American Economic Review*, 2013, *103* (1), 178–219.

Fink, Günther and Felix Masiye, "Assessing the impact of scaling-up bednet coverage through agricultural loan programmes: evidence from a cluster randomised controlled trial in Katete, Zambia," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 2012, *106* (11), 660–667.

Gertler, Paul and Jonathan Gruber, "Insuring Consumption Against Illness," *American Economic Review*, 2002, *92* (1), 51–70.

Isik, Murat and Wanhong Yang, "An analysis of the effects of uncertainty and irreversibility on farmer participation in the Conservation Reserve Program," *Journal of*

Agricultural and Resource Economics, 2004, pp. 242–259.

Jack, B., "Private information and the allocation of land use subsidies in Malawi," *American Economic Journal: Applied Economics*, 2013, *forthcoming.*

Jayachandran, Seema, "Liquidity Constraints and Deforestation: The Limitations of Payments for Ecosystem Services," *American Economic Review*, 2013, *103* (3), 309–13.

Karlan, Dean and Jonathan Zinman, "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment," *Econometrica*, 2009, *77* (6), 1993–2008.

MacKenzie, Ian A, Markus Ohndorf, and Charles Palmer, "Enforcement-proof contracts with moral hazard in precaution: ensuring 'permanence' in carbon sequestration," *Oxford economic papers*, 2012, *64* (2), 350–374.

Mahajan, Aprajit and Alessandro Tarozzi, "Time Inconsistency, Expectations and Technology Adoption: The case of Insecticide Treated Nets," *Working Paper*, 2011.

Mason, Charles F and Andrew J Plantinga, "The additionality problem with offsets: Optimal contracts for carbon sequestration in forests," *Journal of Environmental Economics and Management*, 2013.

Pattanayak, Subhrendu K., Sven Wunder, and Paul J. Ferraro, "Show me the money: Do payments supply environmental services in developing countries?," *Review of Environmental Economics and Policy*, 2010, *4* (2), 254.

Pfaff, Alexander, Juan Andres Robalino, and G Arturo Sanchez-Azofeifa, "Payments for Environmental Services: Empirical analysis for Costa Rica," 2008.

Rust, John, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, 1987, pp. 999–1033.

Saka, A.R., W.T. Bunderson, O.A. Itimu, H.S.K. Phombeya, and Y. Mbekeani, "The effects of Acacia albida on soils and maize grain yields under smallholder farm conditions in Malawi," *Forest Ecology and Management*, 1994, *64*, 217–230.

Salas, Paula Cordero, Brian E Roe, and Brent Sohngen, "Addressing additionality in REDD contracts when formal enforcement is absent," *World Bank Policy Research Working Paper*, 2013, (6502).

Schatzki, Todd, "Options, uncertainty and sunk costs: An empirical analysis of land use change," *Journal of Environmental Economics and Management*, 2003, *46* (1), 86–105.

Suri, Tavneet, "Selection and comparative advantage in technology adoption," *Econometrica*, 2011, *79* (1), 159–209.

Zomer, R., A. Trabucco, L. Verchot, and B. Muys, "Land Area Eligible for Af-

forestation and Reforestation within the Clean Development Mechanism," *Mitigation and Adaptation Strategies for Global Change*, 2008, *13*, 219–239.

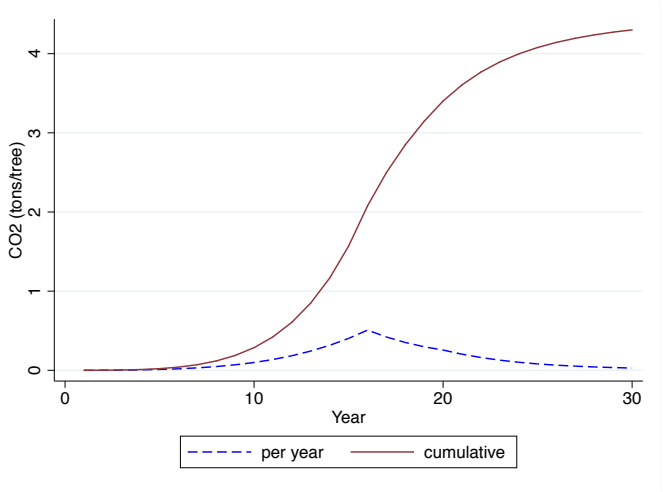


Figure 1: Carbon sequestration associated with *Faidherbia*

Notes: Annual and cumulative CO_2 sequestration per tree over a 30 year period. Sequestration rates are based on allometric equations adapted for the biomass and growth patterns of *Faidherbia albida*.

		Group level: Input subsidy (A) (Cost recovery price = 12,000 ZMK)			
		A = 12,000 (full subsidy)	A = 8,000	A = 4,000	A = 0 (full price)
Individual level: Take up timing	Draw reward then elicit take up	Individual level: Randomly vary R (continuous) between 0 and 150,000 ZMK			
	Elicit take up then draw reward (surprise reward treatment)				

Figure 2: Experimental design



Figure 3: Histogram of tree survival at zero and positive reward levels

Notes: Histogram of tree survival outcomes, by zero and positive reward levels. Tree survival outcomes are conditional on take up. The vertical dashed line shows the performance threshold for the reward.

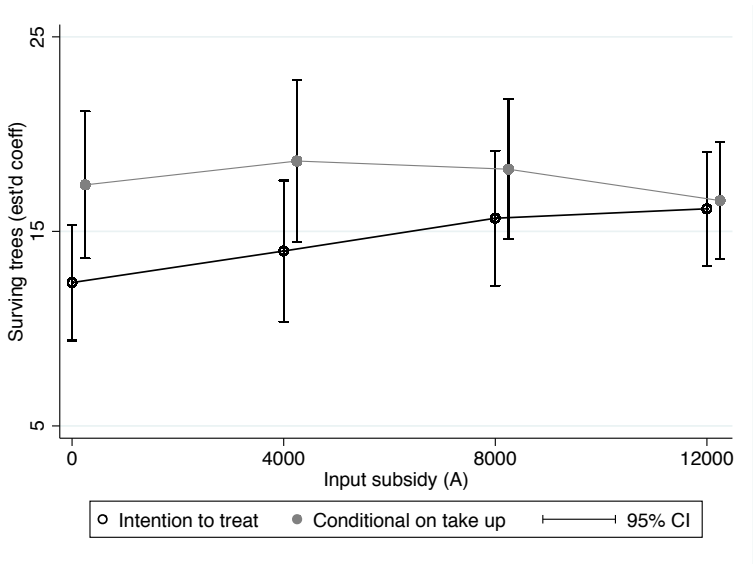


Figure 4: Effect of self-selection on tree survival, by input subsidy treatment

Notes: Coefficients and clustered standard errors from an OLS regression of tree survival on input subsidy categories (ZMK), controlling for reward, reward timing and monitoring treatments. Circles show estimated coefficients for all farmers, replacing outcomes with zeros for non-participants. The gray lines hows estimated coefficients conditional on take up, and represents the test for selection on input subsidy level.

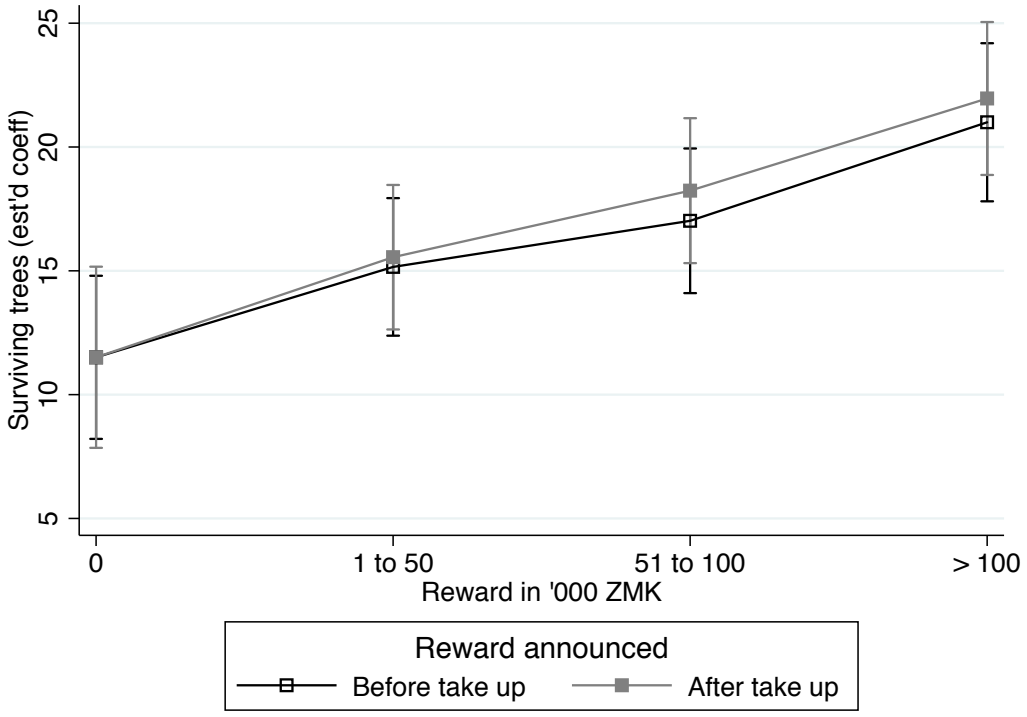


Figure 5: Effect of self-selection on survival, by reward timing treatment

Notes: Figure plots estimated coefficients from an OLS regression tree survival on reward categories interacted with the surprise reward treatment, conditional on take up. The black line plots coefficients for farmers who were informed of their reward before take up. The gray line plots coefficients for farmers unaware of the reward at the time of take up (surprise reward treatment). Error bars show the 95 percent confidence interval.

Table 1: Effect of treatments on contract take up

Dependent variable is program take up	(1)	(2)	(3)	(4)
Input subsidy (A) = 4000	0.0481	-0.0033	-0.0771	
	[0.0768]	[0.0802]	[0.1332]	
Input subsidy (A) = 8000	0.1535**	0.1521**	0.2330**	
	[0.0657]	[0.0699]	[0.1092]	
Input subsidy (A) = 12000	0.2628***	0.2503***	0.3511***	
	[0.0571]	[0.0579]	[0.0865]	
Reward ('000 ZMK)		0.0007**	0.0012*	0.0005*
		[0.0003]	[0.0007]	[0.0003]
A=4000 X Reward			0.0009	
			[0.0010]	
A=8000 X Reward			-0.0012	
			[0.0009]	
A=12000 X Reward			-0.0015**	
			[0.0007]	
Group Fixed Effects	No	No	No	Yes
Individual Controls	No	No	No	No
# obs	1317	624	624	624
Dependent variable mean	0.8292	0.8397	0.8397	0.8397

Notes: OLS regressions of the binary take up outcome on program treatments, with standard errors clustered at the group level. Column 1 includes all participants who made a take up decision. Columns 2-4 exclude participants in the surprise reward treatment. * p<0.10 ** p<0.05 *** p<0.01.

Table 2: Effect of rewards on tree survival

	# trees		1.(Zero trees)		1.(trees ≥ threshold)		# trees > 35	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A. Conditional on take up, surprise reward treatment only							
Reward in '000 ZMK	0.0715***		-0.0017***		0.0013***		0.0052	
	[0.0152]		[0.0004]		[0.0004]		[0.0083]	
Reward > 0		7.2741***		-0.1951***		0.1178**		-0.1129
		[1.8429]		[0.0601]		[0.0457]		[1.0501]
Dep var mean, R = 0	11.3182		0.5303		0.1515		43.0000	
N	568		568		568		128	
	Panel B. All farmers, ITT							
Reward in '000 ZMK	0.0648***		-0.0015***		0.0012***		0.0010	
	[0.0107]		[0.0003]		[0.0003]		[0.0063]	
Reward > 0		6.6610***		-0.1428***		0.1256***		-0.3216
		[1.2943]		[0.0432]		[0.0324]		[1.0394]
Dep var mean, R = 0	10.0294		0.5441		0.1176		43.2857	
N	1317		1317		1317		246	

Notes: OLS regressions of tree survival outcomes on reward conditions. Standard errors in brackets are clustered at the group level. Panel A shows results for self-selected farmers surprise reward treatment), conditional on take up. Panel B shows all farmers, with tree survival outcomes equal to zero for non-participants. All regressions include controls for input cost treatment and monitoring condition. Dependent variables are the continuous number of surviving trees (1 and 2), an indicator for zero trees (3 and 4), an indicator for tree survival at or above the threshold (5 and 6), and the continuous number of surviving trees conditional on being above the threshold (7 and 8). * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Appendices

Table A.1: Balance

	A=0 Mean [SD]	A > 0	R=0 Mean [SD]	Reward > 0	Surprise=0 Mean [SD]	Surprise reward	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Respondent is head of household	0.735	0.001	0.694	-0.0002	0.702	0.038	1292
	[0.442]	[0.003]	[0.463]	[0.0002]	[0.458]	[0.025]	
Age, respondent	37.872	-0.074	37.439	0.0043	37.25	1.284*	1266
	[13.716]	[0.100]	[12.808]	[0.0065]	[13.684]	[0.709]	
Female headed household	0.135	0.001	0.149	0.0001	0.119	0.009	1292
	[0.343]	[0.003]	[0.358]	[0.0002]	[0.324]	[0.017]	
Years of education, respondent	5.342	-0.037	5.284	0.0019	5.373	-0.011	1289
	[3.276]	[0.027]	[3.133]	[0.0019]	[3.329]	[0.176]	
Household size	5.465	-0.036**	5.328	-0.0002	5.246	0.208*	1292
	[2.142]	[0.015]	[2.390]	[0.0012]	[2.217]	[0.114]	
Ordinal discount rate (1 - 5)	2.454	-0.001	2.538	0.0013	2.43	0.106	1262
	[1.627]	[0.013]	[1.714]	[0.0009]	[1.612]	[0.095]	
Non-agricultural assets	9.806	-0.101**	9.343	-0.0027	9.08	0.314	1292
	[5.789]	[0.041]	[5.625]	[0.0030]	[5.506]	[0.287]	
Years working with Dunavant	4.255	-0.032	3.869	-0.0022	3.904	0.053	1280
	[3.754]	[0.036]	[3.406]	[0.0021]	[3.500]	[0.211]	
Total landholdings (hectares)	3.027	-0.023	2.881	-0.001	2.874	0.057	1287
	[2.256]	[0.023]	[2.188]	[0.0012]	[2.256]	[0.115]	
Number of fields	2.874	0.001	2.866	0.000	2.886	-0.054	1292
	[1.065]	[0.009]	[1.194]	[0.0007]	[1.122]	[0.070]	
Average distance from home to plots	20.505	-0.087	19.628	-0.0069	18.261	0.909	1271
	[24.221]	[0.235]	[22.539]	[0.0130]	[20.650]	[1.080]	
Poor soil fertility	0.108	-0.001	0.104	0.0001	0.106	-0.030*	1290
	[0.311]	[0.002]	[0.307]	[0.0002]	[0.308]	[0.018]	
YGL has visited respondent's land	0.704	0.001	0.654	0.0001	0.73	-0.027	1249
	[0.457]	[0.003]	[0.478]	[0.0003]	[0.444]	[0.024]	
Affiliated with CFU or COMACO	0.037	0.002	0.037	0.000	0.042	0.003	1292
	[0.189]	[0.002]	[0.190]	[0.0001]	[0.202]	[0.012]	
Prior knowledge of Faidherbia	0.68	-0.001	0.664	0.000	0.64	0.029	1290
	[0.467]	[0.004]	[0.474]	[0.0003]	[0.480]	[0.025]	
Prior planting of Faidherbia	0.111	-0.001	0.09	0.0002	0.088	0.014	1289
	[0.314]	[0.002]	[0.287]	[0.0001]	[0.284]	[0.014]	
Knowledge of risks to tree survival	1.72	-0.005	1.701	-0.0003	1.648	-0.013	1292
	[0.905]	[0.006]	[0.785]	[0.0004]	[0.816]	[0.046]	
N	325	967	134	1158	614	678	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: Means are reported for the base group in columns 1, 3 and 5. Coefficients and standard deviations from a regression of the household variable on treatment are reported in other columns.

Table A.2: Attrition across data collection phases

	Takeup Mean [SD]	Baseline	Endline	Tree monitoring
	(1)	(2)	(3)	(4)
Input cost	6156.4161	0.0000*	0.0000	0.0000
	[4539.9090]	[0.0000]	[0.0000]	[0.0000]
Reward ('000 ZMK)	76.9909	-0.0002	0.0000	0.0000
	[51.8237]	[0.0001]	[0.0001]	[0.0000]
Surprise reward treatment	0.5239	0.0097	0.0108	0.0025
	[0.4996]	[0.0088]	[0.0123]	[0.0061]
N, outcome = 1	1317	1292	1232	1083

Notes: Attrition across data collection rounds by treatment. Column 1 shows means and standard deviations for each treatment. Each cell in columns 2 - 4 shows the coefficient from a regression of an indicator being present at the data collection stage regressed on each treatment with standard errors clustered at the farmer group level. Column 4 is conditional on participating in the program (N=1092). * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table A.3: Procrastination

Dependent variable:	Takeup	Survival	Survival	Takeup	Takeup
	(1)	(2)	(3)	(4)	(5)
	Panel A: Self-described procrastinator				
Binary Procrastination Measure	-0.0480**	-0.6979	-1.0575	-0.1028**	-0.0960
	[0.0235]	[1.2464]	[1.2254]	[0.0470]	[0.0663]
Input subsidy in '000 ZMK	0.0210***	-0.0477	-0.0108	0.0177***	0.0179***
	[0.0043]	[0.1960]	[0.1913]	[0.0043]	[0.0054]
Reward in '000 ZMK			0.0717***		0.0005
			[0.0122]		[0.0003]
Procrastination x Input subsidy				0.0089*	0.0054
				[0.0049]	[0.0071]
Constant	0.4942***	6.6855*	1.8827	0.5213***	0.4800***
	[0.0882]	[3.4530]	[3.4798]	[0.0888]	[0.1196]
N	1157	978	978	1157	544
	Panel B: Reports procrastination on other activities				
Binary Procrastination Measure	-0.0221	1.0007	1.0920	-0.0851	-0.0056
	[0.0293]	[1.3377]	[1.3321]	[0.0676]	[0.0767]
Input subsidy in '000 ZMK	0.0212***	-0.0768	-0.0421	0.0180***	0.0188***
	[0.0043]	[0.1975]	[0.1934]	[0.0045]	[0.0056]
Reward in '000 ZMK			0.0706***		0.0005
			[0.0122]		[0.0003]
Procrastination x Input subsidy				0.0101	0.0055
				[0.0071]	[0.0080]
Constant	0.4746***	6.7166**	1.8782	0.4924***	0.4303***
	[0.0870]	[3.3847]	[3.4003]	[0.0881]	[0.1123]
N	1165	984	984	1165	549

Notes: OLS regressions of take up and survival on indicators of procrastination. Standard errors clustered at the group level are in brackets. Columns 2 and 3 condition on take up. Column 5 conditions on knowing the reward before take up (excludes the surprise reward treatment. See text for a description of the procrastination measures used in the regressions. * p<0.10 ** p<0.05 ***