Workshop / *Atelier*

Learning for adopting

Apprendre pour adopter

Wednesday-Thursday, June 1st-2nd, 2016
Mercredi-Jeudi 1er-2 Juin 2016

Technology adoption under uncertainty: Take up and subsequent investment in Zambia

by

B. Kelsey Jack, Tufts University and NBER

Paulina Oliva, UCSB and NBER

Samuel Bell, Shared Value Africa

Christopher Severen, UCSB

Elizabeth Walker, NERA Consulting

Technology adoption under uncertainty: Take up and subsequent investment in Zambia*

B. Kelsey Jack
Tufts University and NBER

Paulina Oliva
UCSB and NBER

Samuel Bell
Shared Value Africa

Christopher Severen
UCSB

Elizabeth Walker
NERA Consulting

Abstract

Technology adoption often requires investments over time. As new information about the costs and benefits of investment is realized, agents may prefer to abandon a technology that appeared profitable at the time of take-up. This re-optimization can reduce the cost-effectiveness of adoption subsidies. We use a field experiment with two stages of randomization to generate exogenous variation in the payoffs associated with take-up and subsequent investment in a new technology: a tree species that provides private fertilizer benefits to adopting farmers. Our empirical results show high rates of abandoning the technology, even after paying a positive price to take it up. The experimental variation offers a novel source of identification for a structural model of intertemporal decision making under uncertainty. Estimation results indicate that the farmers experience idiosyncratic shocks to net payoffs after take-up, which increase take-up but lower average per farmer tree survival. We simulate counterfactual outcomes under different levels of uncertainty and observe that farmers with high returns are able to self-select at take-up only when the level of uncertainty is relatively low. Thus, uncertainty provides an additional explanation for why many subsidized technologies may not be utilized even when take-up is high.

*Helpful comments were received from Jim Berry, Chris Costello, Andrew Foster, Alex Pfaff, Andrew Plantinga, Stephen Ryan, Kenneth Train, Mushfiq Mobarak, Ryan Kellogg, Tavneet Suri and audiences at numerous seminars and conferences. The authors thank the IGC, CDKN and Musika for financial support, and the Center for Scientific Computing at the CNSI and MRL at UC Santa Barbara (NSF MRSEC DMR-1121053 and NSF CNS-0960316) for use of its computing cluster. Field work was facilitated by Innovations for Poverty Action, with specific thanks to Jonathan Green, Farinoz Daneshpay, Mwela Namonje and Monica Banda. The project was made possible by the collaboration and support of Shared Value Africa and Dunavant Cotton, Ltd.

1 Introduction

Many technology adoption decisions –in development, health and environmental policy– consist of at least two parts, which occur at different points in time: an initial take-up decision and a subsequent investment or follow-through decision. While subsidies are often used to increase take-up, critics of subsidies for technology adoption worry that subsidizing the initial take-up decision may lower subsequent follow-through, leading to a misallocation of subsidized technologies to adopters who never use them.

The empirical evidence on whether the take-up price is correlated with follow-through is mixed: some studies find a positive correlation (e.g., Ashraf et al. 2010) and others none (e.g., Cohen and Dupas 2010).[1] The literature has put forward many reasons why follow-through may or may not be affected by the initial cost of the technology, including screening effects, learning, and psychological channels such as sunk costs or procrastination (Ashraf et al. 2010; Cohen and Dupas 2010; Mahajan and Tarozzi 2011; Ashraf et al. 2013; Dupas 2014; Beaman et al. 2014; Fischer et al. 2014; Carter et al. 2014; Cohen et al. 2015). With the exception of learning, little attention has been paid to the role of dynamics and uncertainty in the initial take-up decision.[2] Specifically, at the time of take-up, many of the benefits and costs associated with the follow-through decision may be unknown. New information may arrive after take-up in the form of learning about the technology (Foster and Rosenzweig 1995; Conley and Udry 2010) or in the form of transient shocks to the opportunity cost of follow-through. If the new information is bad news about the profitability of the technology, then adopters may opt for abandoning the technology. Adopters know that they can reoptimize once new information is available and are likely to account for this at the time of take-up. Thus, the take-up decision can be interpreted as the purchase of an option to follow-through.

[1]Similar issues arise in cost sharing of medical treatment (e.g., Goldman et al. (2007)).

[2]The dynamic effects of subsidies on subsequent demand for the same technology is investigated by Carter et al. (2014), Dupas (2014) and Fischer et al. (2014), all of which find support for learning.

This paper identifies theoretically and empirically the role of uncertainty in the decision to take-up and follow-through with a new technology in a setting where take-up may be subsidized. We apply a dynamic conceptual model to farmer decisions to adopt a type of fertilizer tree in Zambia. This technology requires both a one-time take-up decision (purchasing seedlings) and a follow-through decision (planting and caring for the trees). Farmers face potential shocks to the opportunity cost of following-through with the trees, including illness of a household member, pests, drought or other factors that may affect (either positive or negative) other crops and/or the trees, and other events in the household or on the farm that may be hard to define or measure. Our approach to identifying uncertainty in our field setting circumvents the challenge of comprehensive measurement of all components of opportunity cost, including those that are not realized, by extending a revealed preference framework to the dynamic case: adoption-related choices by the same individual on the same technology are made at different points in time and reveal the information that was acquired in between. Our study proceeds in four steps: (1) a conceptual model of technology adoption under uncertainty; (2) a field experiment with variation in adoption payoffs at two points in time; (3) a structural model that builds on (1) and (2); and (4) counterfactual simulations that vary the magnitude of uncertainty and show implications for adoption outcomes.

To develop intuition, we begin with a stylized model of intertemporal adoption under uncertainty in the presence of subsidies, where individuals make binary take-up and follow-through decisions at two different points in time.[3] Between these two points in time new information about the opportunity cost of follow-through is acquired. The theoretical model generates clear predictions about the relationship between uncertainty and adoption outcomes. First, a mean-preserving increase in uncertainty makes take-up more attractive provided that abandoning the technology at a later stage is costless. This is because, re-

[3] Our empirical model adds an intensive margin to the follow-through decision, similar to Ashraf et al. 2010; Cohen and Dupas 2010; Fischer et al. 2014 and others.

gardless of how costly following through turns out to be, profit is always bounded below at zero by the option to abandon the technology. Thus, uncertainty can only increase the upside of the take-up decision.[4] Second, uncertainty undermines the screening effect of the take-up price. Intuitively, if adopters know little about their net cost of follow-through when they take-up, then a higher take-up price will not be effective at screening out those who will end up having a high cost at follow-through. Our conceptual framework borrows heavily from the literature on investment under uncertainty (Pindyck 1993; Dixit and Pindyck 1994) and, like Fafchamps (1993), shows that choices that appear to lead to losses (like purchasing a technology that is soon to be abandoned) can be rational ex-ante if their purpose is to preserve flexibility.[5]

Next, we use a multi-period field experiment in rural Zambia to generate empirical evidence for the presence of uncertainty. Farmers choose whether to adopt a tree species that generates private soil fertility benefits over the long term, but carries short-run costs.[6] We observe whether the 1,314 farmers in the study take up a 50-tree seedling package at the start of the agricultural cycle. The follow-through decision consists of the number of seedlings that the farmer chooses to plant and care for (which we together refer to as tree cultivation) and occurs over the course of the subsequent year. Shocks to the opportunity cost of follow-through may cause farmers to abandon the technology, which they can do without penalty. We measure follow-through as tree survival after one year, and assume that farmers can guarantee tree survival for some level of costly effort.[7]

[4] This is true even in the presence of insurance, as the costless exit is still present in this case and thus the contract becomes a substitute for insurance. See Giné and Yang (2009) for an example of how an uninsured credit contract may be more attractive than an insured one in the presence of limited liability.

[5] Other applications of dynamic decision making under uncertainty in the development and environmental literature include Bryan et al. (2014); Magnan et al. (2011); Arrow and Fisher (1974).

[6] Positive externalities, such as carbon sequestration and reduced soil erosion, further justify the subsidy from a policy perspective.

[7] The choice of minimum effort that guarantees survival is optimal under convexity of the survival risk function as a function of effort. The only source of uncertainty in tree survival in our model is the farmer's endogenous choice of effort in response to new information about the costs of follow-through. This assumption is examined in greater detail in Appendix A.3.

We introduce exogenous variation into this adoption decision at two different points in time. First, we vary the take-up cost through a subsidy on the purchase of a seedling package. Farmers' response to this random variation helps characterize the heterogeneity in expected costs across farmers.[8] Second, we vary the payoff to follow-through by varying the size of a reward that is conditional on the survival of at least 35 trees one year after take-up. The tree cultivation choices farmers make in response to the reward help us characterize the distribution of follow-through costs after potential shocks have been realized. Under the assumption that shocks are independent across farmers, the difference in the variance of net costs between the two points in time can be attributed to uncertainty.[9] Note that, rather than artificially varying the allocation of shocks across our study population, the reward creates exogenous variation in the variance of possible outcomes faced by the farmers, and therefore in the distribution of shocks. By offering a positive payoff, the performance reward varies the distribution of shocks much in the way that varying the terms of an insurance contract has a state-contingent effect on the distribution of outcomes.[10] Together, the different sources of variation at two points in time identify a structural model of intertemporal decisions that can distinguish between static and dynamic explanations for the outcomes that we observe. This is among the first papers to introduce multiple dimensions to the experimental design to distinguish between adoption decisions and returns to investment (see Karlan and Zinman (2009) and extensions of their design by Ashraf et al. (2010); Cohen and Dupas (2010) and others), and the first to use this research design to explore time-varying returns to investment.

The reduced form responses to the randomized treatments are broadly consistent with the predictions of our theoretical model in the presence of uncertainty: while farmers respond

[8] Liquidity constraints could also affect the decision to take-up. To minimize the importance of cash-on-hand, farmers receive a show-up fee sufficient to cover take-up costs. We also test for self-selection based on broader forms of liquidity constraints: using random variation in the timing of the reward announcement. We discuss these tests for liquidity constraints and other confounds in Section 4 and Appendix A.4.

[9] The cross-farmer independence assumption rules out common shocks. In a model variant, discussed in Section 5, we relax the independence assumption by allowing for an unexpected common shock to all farmers.

[10] A similar approach is used by Einav et al. (2013); Bryan et al. (2014); Karlan et al. (2014), among others.

to economic incentives (they take-up at higher rates under higher subsidies and follow-through at higher rates under higher rewards), the price at which each individual takes up is not predictive of the follow-through outcome (i.e. we find no significant screening effect of prices).[11] In addition, a large share of farmers who paid a positive price end up abandoning the technology altogether.[12] Although these facts suggest that uncertainty plays a role in farmers' decisions, they do not allow us to quantify the amount of uncertainty farmers face nor how important it is for farmers' decisions versus other forms of heterogeneity that can lead to similar behavior.

We turn next to our structural model to shed further light on the role and magnitude of uncertainty in our setting, and the generalizability of our findings. We start by noting that uncertainty is not the only plausible explanation for the absence of positive screening effects of prices. When follow-through has an intensive margin, there may be heterogeneity in both the level of the profit (for example, if there are fixed costs to adoption) and in the number of trees that maximizes the profit (i.e. the interior solution to the farmers' profit maximization problem); moreover, these two types of heterogeneity may be positively or negatively correlated. Only a positive correlation between the level of private profit and privately optimal number of trees would generate higher follow-through rates among those who participate at higher cost, yet a negative correlation is empirically plausible. Suri (2011), for example, offers evidence of a negative correlation between optimal rates of usage and fixed costs of adoption in the case of hybrid crop varieties. Our structural model allows

[11]The lack of self-selection in our setting stands in contrast with Jack (2013), who provides evidence that farmers self-select based on future costs into a tree planting incentive contract in Malawi. She studies a different context and different contract design. In addition, the pattern of selection effects over time in her study is consistent with a multi-year extension of our conceptual framework, which would predict stronger selection as the number of farmers who continue to cultivate trees shrinks.

[12]We rule out a number of alternative motives for this behavior. First, we rule out side-selling by exploiting cross-group variation in incentives to side-sell. Second, we test whether a desire to please the experimenter (social desirability bias) drives our results by allowing for a common "boost" to the attractiveness of take-up in our structural model, and find it has little effect on our estimates (Section 6). We also examine the effect of time inconsistent preferences (as in Mahajan and Tarozzi 2011) in Section 7 and Appendix table A.5.6.

for heterogeneity in the privately optimal number of trees as well as in the net cost of follow-through, and allows these two known (to the farmer) components of private profit to be freely correlated.[13] We find that in our setting heterogeneity along the intensive margin of tree survival operates in oposite direction to the extensive margin heterogeneity (similar to Suri (2011)), thus weakening the screening effect of prices. In addition, our estimates find a large variance in the unknown component of costs; i.e. a large amount of uncertainty. To illustrate its magnitude, we calculate that 15 percent of farmers would change their ex ante decision about meeting the threshold if they could take the new information into consideration.

Given that both uncertainty and heterogeneity along the intensive margin are contributing to the lack of screening coming from the take-up price, we implement counterfactual simulations to better understand the relative importance of uncertainty in explaining our results. We find that at levels of uncertainty lower than those in our empirical setting, higher prices for take-up do have a positive effect on follow-through. Reducing the variance of shocks by 50 percent (everything else constant) would bring up follow-through rates among those who take-up under full price by 15 percent, because of improved screening at take-up.

Our conceptual framework and simulations also highlight the different role for subsidies in the presence of uncertainty. Greater uncertainty in the payoffs from follow-through makes subsidies less important for take-up because the option value, which increases with uncertainty, drives up the expected profit. However, with high uncertainty, the more modest effect that subsidies may have on take-up may be (almost) free of adverse selection effects, making subsidies less problematic for allocational efficiency.

Methodologically, our econometric framework is an example of sequential identification of subjective and objective opportunity cost components in a dynamic discrete choice model (Heckman and Navarro 2007, 2005). As described in Heckman and Navarro (2007), we can

[13]This is akin to correlated random coefficient (CRC) models, where returns to the technology are allowed to differ across potential adopters and therefore influence their decision to adopt (Heckman et al. 2010).

account for selection into treatment (in our case, take-up) when identifying the distribution of the unobserved opportunity cost determinants. We do so by introducing two layers of random variation in economic incentives, one of which produces a probability of take-up equal to one for a randomly selected sub-population and a second of which produces an interior solution in tree cultivation outcomes with probability one in the limit. The use of experimental variation in treatments at two different points in time offers an alternative to a panel data structure (used for example, in Einav et al. (2013)), since statistically independent samples are exposed to each of the different treatment combinations. To our knowledge, this is the first paper to introduce experimental variation in order to satisfy the exclusion restrictions needed for sequential identification.

The paper proceeds as follows. We begin with a simple theoretical model to generate intuition. Section 3 describes the empirical context and experimental design, and Section 4 shows reduced form results. We present the empirical model and its identification in Section 5 and show estimation results and simulations in Section 6. Section 7 discusses interpretation and Section 8 concludes.

2 A simple model of intertemporal technology adoption

Consider a two period model, where each agent chooses whether to purchase (take-up) a single unit of a technology in the first period (time 0) , and whether to follow-through with implementation of the technology in the second period (time 1). The immediate cost of taking up is $c - A$, where c is the market price of the technology and A is an exogenous subsidy. The benefit of following-through is given by $R - (F_0 + F_1)$, where $F_0 + F_1$ is the "net private cost" of following through and R is an exogenous reward for doing so. Since $F_0 + F_1$ is net of benefits, it can be positive or negative. The first component of the net cost, F_0, is known to the agent at the time of take-up, while F_1 is unknown to the agent at time 0 and its realization (which is revealed to the agent at time 1) has a known distribution that is constant across agents. Note that although F_0 is known at time 0, both F_0 and F_1 are incurred at time 1. Assume that c, A and R are constant across agents, while F_0 varies according to some cdf $G_0(f_0)$. Assume further that

(i) F_0 and F_1 are independent, and

(ii) $\mathbb{E}_{t=0}(F_1) = \mathbb{E}_{t=1}(F_1) = 0$ (i.e. agents have rational expectations).

Under these assumptions, F_0 represents the agent's best guess at $t = 0$ about her specific net cost of following through, and F_1 represents any new information that emerges after the take-up decision is made.

Following backward induction, the agent decides to follow-through at $t = 1$ if $R - F_0 - F_1 > 0$. If this inequality does not hold, the agent receives a payoff of zero at $t = 1$. At $t = 0$, the agent decides to take-up by purchasing the technology if

$$c - A - \delta \mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) < 0 \tag{1}$$

where δ is the one-period discount factor and the expectation in (1) is taken with respect to the density of F_1.

To simplify the exposition, assume that the distribution of F_1 is such that $F_1 \in \{f_L, f_H\}$, with $f_L < f_H$ and $\Pr(F_1 = f_L) = p_L$. Thus we can represent a mean-preserving increase in uncertainty as a symmetric widening of the distance between f_L and f_H.[14] With this assumption, we can classify individuals into three types: those who always follow through, regardless of the realization of F_1 (*always follow-through types*), those who follow through only if the low net cost shock is realized (*contingent follow-through types*), and those who

[14]This model simplifies our empirical setting in two key ways: first, it assumes a binary follow-through decision and second, it assumes a discrete distribution on F_1. As we show when we present our empirical model, the propositions derived from this model are not an artifact of the distributional assumption on F_1 nor of the binary decision that characterizes follow-through in this simple model.

never follow-through (*never follow-through types*). These three types of agents can be characterized by whether their value of F_0 is below $R - f_H$, between $R - f_H$ and $R - f_L$, and above $R - f_L$, respectively. Figure 1 graphically shows the proportions for each type of agent using areas under a symbolic bell-shaped distribution for F_0, separated by gray dashed lines. Figure 1 also illustrates two thresholds (along the support of F_0) for take-up in black dashed lines. The first take-up threshold (labeled $R - \mathbb{E}(F_1) - \frac{c-A}{\delta}$) is only binding if it falls to the left of the threshold that defines always adopters ($R - f_H$). When this first take-up threshold binds, only always follow-through types take-up. The second take-up threshold (labeled $R - f_L - \frac{(c-A)}{\delta p_L}$) is perhaps more interesting. When binding, all always-follow through types take-up, but only a share of contingent follow-through types take-up (those to the left of the threshold). We use this figure to explain intuitively the results outlined by each of our propositions below. The formal proofs of these propositions can be found in Appendix A.1.

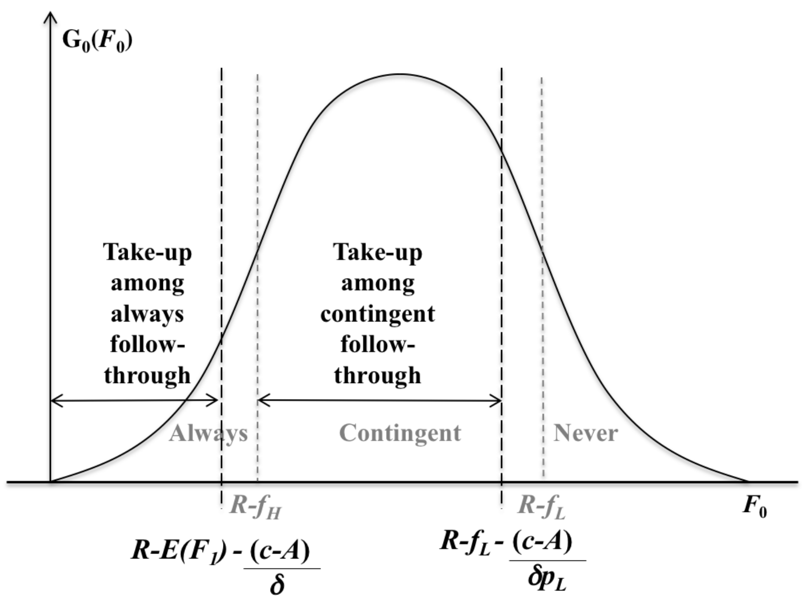


Figure 1: Take-up and follow-through thresholds as a function of agent type

Notes: The figure shows the shares of always adopters, contingent adopters and non-adopters over a symbolic probability density function of F_0. The grey thresholds ($R - f_H$ and $R - f_L$) correspond to the follow-through thresholds, while the black thresholds correspond to the take-up thresholds.

Proposition 1 *Follow-through conditional on take-up increases as a function of take-up cost, i.e. there is a screening effect of the take-up cost.*

To see this, note that as take-up cost increases (represented by $c - A$ in Figure 1), the second take-up threshold moves to the left, bringing down the overall share of contingent follow-through types among the set of individuals who take-up. Since contingent adopters follow-through with probability less than one (p_L), this in turn increases the share of individuals who follow-through among those who take-up.[15]

Proposition 2 *An increase in uncertainty reduces follow-through conditional on take-up.*

This can also be appreciated from Figure 1: a widening of the distance between f_L and f_H causes the share of contingent follow-through types to increase (as the two grey dashed lines move further apart). Note that as uncertainty increases, the position of the second take-up threshold does not change relative to the threshold that determines the upper bound for contingent follow-through types. Thus, this group becomes a larger share of those who take up, reducing average follow-through.

Corollary 2.1 *Under no uncertainty, everyone who takes-up follows-through.*

This is easy to see from Figure 1: under no uncertainty (where $f_L = f_H$) there would be only always follow-through types and never follow-through types.

Proposition 3 *An increase in uncertainty weakens the relationship between take-up cost and conditional follow-through shown in Proposition 1.*

To see this, consider the takeaways of Propositions 1 and 2 simultaneously. The share of contingent adopters that are excluded by an increase in the take-up cost becomes a smaller

[15]If the take-up cost, $c - A$, increases enough that the first take-up threshold is binding, follow-through conditional on take-up reaches 100 percent and is constant for further increases in the take-up cost.

proportion of all those who take-up when uncertainty increases.

Proposition 4 *The option value associated with take-up is increasing in uncertainty, which results in higher take-up at all take-up cost levels.*

This is shown formally in the appendix along with the formal definition of option value in our context. Intuitively, the option value is the value of reoptimizing once new information (the realization of F_1) emerges. As the distance between f_H and f_L increases, the payoff at $t = 1$ conditional on a low cost shock (f_L) increases for contingent follow-through types. Because agents can choose not to follow through, the payoff at $t = 1$ conditional on a high cost (f_H) stays constant at zero. Thus, the expected value of the contract at $t = 0$ increases with uncertainty, and this increase emerges solely because of the possibility of reoptimizing (i.e. choosing not to follow-through). This results in higher take-up.

A note on risk neutrality. We assume linear utility – or risk neutrality – throughout the paper, including the empirical analysis. Assuming some degree of risk aversion would not change our results qualitatively, although it would lower the value placed on extreme positive profitability shocks at the time of take-up. This would make the expected value of the contract at $t = 0$, and therefore take-up, less responsive to increases in uncertainty. That said, the risk neutrality assumption is relatively innocuous and carries important advantages given our empirical context. Although risk aversion is an important component of intertemporal decisions with costs or benefits that represent substantial shares of household income, our specific technology adoption decision causes relatively small changes to income. In addition, our framework (both theoretical and empirical) models decisions as a function of the profits associated with adoption relative to the best alternative use of household resources. Thus, a positive shock to the opportunity cost of adoption could correspond to an increase or a decrease in overall household income. For example, an increase in profitability of a competing economic activity and a labor shortage due to health could both represent an increase in the opportunity cost of adoption, but would have opposite effects on total income and thus on marginal utility of income. Incorporating risk aversion into our theoretical model would require us to make modeling assumptions about the nature of the opportunity cost of adoption. Hence, assuming risk neutrality allows us to leave the source of the opportunity cost unspecified, which makes our framework generalizable to any source of uncertainty regardless of its impact on overall income.

Transitory shocks and learning. So far, we have left open the question of whether F_1 should be interpreted as a persistent or a transitory shock and our framework is consistent with both interpretations. However, the distinction matters for future take-up decisions. If the F_1 component of the returns to the technology is persistent, future take-up decisions will occur under a lower level of uncertainty. If F_1 is transitory, future take-up decisions will look similar to the first take-up decision. We cannot completely disentangle these two interpretations of the model in our context, though we use survey data to provide suggestive evidence on the extent of learning (see Section 7).

3 Context and experimental design

We bring the propositions from our conceptual model to a two-part technology adoption problem, characterized by uncertainty in the costs and benefits of following through with the technology. In the context of an ongoing project to encourage the adoption of agroforestry trees (*Faidherbia albida*), we introduce exogenous variation in the payoffs to farmers at the time of their take-up and follow-through decisions. We use the experimental variation to uncover the existing levels of static heterogeneity and uncertainty in the population of farmers, which we model as random parameters. This section describes the context and the experimental design in detail.

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 farmers growing cotton under contract with Dunavant, alongside other subsistence crops. The project is part of the NGO partner's portfolio of carbon market development projects in Zambia.

3.1 The technology

Faidherbia albida is an agroforestry species endemic to Zambia that fixes nitrogen, a limiting nutrient in agricultural production, in its roots and leaves. 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 cropping season, means that planting *Faidherbia* does not displace other crop production (Akinnifesi et al. 2010). Agronomic studies suggest significant yield gains from *Faidherbia*.[16] However, these private benefits take 7-10 years to reach their full value, and may be insufficient to justify the up-front investment costs, particularly if farmers have high discount rates. We observe low adoption rates at baseline: less than 10 percent of the study households reported any *Faidherbia* on their land. This could be explained by low perceived private net-benefits, by high costs associated with accessing inputs – there is no existing market for *Faidherbia* seedlings – or cultivating the trees, or by a lack of information.[17]

Subsidies may therefore be necessary to increase take-up rates, and are justified by positive environmental externalities and market failures that contribute to high private discount

[16]Estimates of yield increases range from 100 to 400 percent, relative to production without fertilizer (Saka et al. 1994; Barnes and Fagg 2003). The benefits relative to optimal fertilizer application are less well understood, but 30 percent of farmers in our baseline survey do not use any fertilizer and those who do use it primarily for cash crops.

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

rates. Environmental benefits include erosion control, wind breaks, and carbon sequestration. Based on allometric equations from Brown (1997), adapted to the growth curves for *Faidherbia*, we estimate that over 30 years, a tree sequesters around 4 tons of carbon dioxide equivalent. Discounting the annual sequestration at 15 percent leads to a present value of around 0.48 tons per tree.

Both the private and the public benefits associated with adoption require that farmers continue to invest in the technology after the initial take-up decision. To keep trees alive, farmers must plant, water, weed and otherwise care for the trees, activities that are costly in the short run. In addition, the opportunity cost of these investments may depend on shocks to household labor supply, weather, pests and prices, all of which are realized after take-up. Therefore, the technology maps clearly onto our conceptual framework.

3.2 Experimental design and data collection

The field experiment was implemented between November 2011 and December 2012 with 125 farmer groups and 1,314 farmers. Implementation of the study relied on Dunavant's outgrower infrastructure, which is organized around sheds, each of which serves several dozen farmer groups. Each farmer group consists 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. Implementation was concentrated at two points in the agricultural season, as shown in Appendix figure A.5.1. First, farmer training, program enrollment, and a baseline survey all occurred at the beginning of the planting season. As the figure shows, this is also the time that farmers make decisions on other crops and technologies. Second, the endline survey, tree survival monitoring and reward payment occurred at the end of our study period, one year after program enrollment. In addition to these main stages, we performed mid-year tree monitoring for a subsample of our farmers and a brief survey at the end of the planting season.

At the training, farmers were provided with instructions on planting and caring for the trees, information about the private fertilizer benefits and public environmental benefits of the trees, and details on eligibility for the program.[18] All farmers who attended the training received a show up fee of 12,000 ZMK and lunch. Farmers were told the money received, which was equivalent to about a day's agricultural wages, was compensation for their time and was theirs to keep. This design feature was intended to reduce the effect of immediate liquidity constraints on take-up.

Enrollment occurred at the end of the training and consisted of farmers' take-up decision. Study enumerators explained the details of the enrollment choice: a take it or leave it offer of 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. The study design varied two major margins of the farmer's decision to adopt *Faidherbia albida*. First, the size of the take-up subsidy (A) varied between 0, 4,000, 8,000, and 12,000 ZMK. At zero subsidy, farmers paid 12,000 ZMK (approximately USD 2.60) for inputs, which is the cost recovery price for the implementing organization, but is likely to fall below farmers' full cost of accessing seeds or seedlings outside of the program. Groups were randomly assigned to one of four take-up subsidy treatments with equal probability using the min max T approach (Bruhn and McKenzie 2009), balanced on Dunavant shed, farmer group size and day of the training. The subsidized price of the inputs was announced to all farmers in the group at the end of training, before the take-up decision was made.

Second, the program offered a threshold payment conditional on follow-through (tree survival) after one year. The payment varied randomly across farmers. Farmers received the reward if they kept 70 percent (35) of the trees alive through the first dry season (for 1

[18]Eligibility required that land must have been un-forested for 20 years, must be owned by the farmer, and must not be under flood irrigation.

year). The threshold reward, as opposed to a per-tree incentive, allows us to draw a sharper distinction between internal incentives and external incentives to cultivate the trees, which aids identification of the structural model. To implement the individual-level randomization of the rewards and allow participants to make their take-up decision in private, the study enumerators called the farmers aside one by one and described the threshold nature of the reward. The farmer then drew a scratch off card from a bucket, which revealed the individual reward value, after which the take-up decision was recorded. The size of the threshold performance reward (R) was varied in increments of 1,000 ZMK, ranging from zero to 150,000 ZMK or approximately 30 USD.[19] Variation in the reward was introduced using a random draw at the time of the take-up decision. One-fifth of all draws were for zero ZMK with the remaining four-fifths distributed uniformly over the range. The frequency of treatment outcomes are shown in Appendix figure A.5.2.

We introduced an additional source of variation that allowed us to test for liquidity constraints as a driver of selection outcomes: the timing of the reward draw was varied at the individual level to occur either before or after the farmer's take-up decision, with 52.5% assigned to the surprise reward treatment. When the reward is known before take-up, it affects both the type of farmer who takes-up and also the decision to follow-through; when it is not known at take-up, it affects only follow-through. Varying when the reward was revealed allows us to isolate its effect on selection, in a similar spirit to Karlan and Zinman (2009).[20]

Following the take-up decision, all farmers were given a baseline survey that lasted for approximately one hour. After the survey, participating farmers signed a contract indicating

[19] 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.

[20] We do not manipulate or measure beliefs about potential financial benefits from joining in the surprise reward treatment, and cannot therefore assume that farmers in the surprise reward treatment assumed $R = 0$ at the time of take-up.

their agreement with the program terms, paid the take-up cost and collected their seedlings. To minimize the effect of seedling quality on tree survival, farmers were not allowed to pick their seedlings.

One year after the training, all farmers in the study sample were given an endline survey. Approximately one week after the endline survey, farmers with contracts were visited for field monitoring, during which the farmer and a study enumerator examined each tree, and recorded whether it was sick, healthy or dead. Monitors also recorded indicators of activities likely to affect survival outcomes: weeding, watering, constructing fire breaks, and field burning (which, in contrast to the other three, threatens tree survival). All surviving trees counted toward the tree survival threshold. Within a couple of days of the monitoring visit, farmers with 35 or more surviving trees received their reward payment. Keeping the payments separate from the monitoring was intended to improve monitors' objectivity.[21]

In addition to the baseline and endline surveys, one-fifth of the farmers were randomly sampled for ongoing data collection on activities and inputs associated with the trees and with other crops. Farmers selected for this effort monitoring received a very short survey (around 20 minutes) every two weeks, during which a project monitor asked the farmer about agricultural activities, including those related to the trees, since the last visit. No information was provided to the farmers about their performance and monitors were instructed not to prompt specific activities or answer technical questions. We control for the effort monitoring subsample in our analysis. The resulting data yield two important facts about the timing of farmer investments. First, planting activities began immediately after the training for some farmers, while other farmers chose to delay tree planting until other crops were planted and the rainfall patterns were clearly established. Second, tree care activities spanned the

[21]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. Given differences in career concerns across monitors (some had higher paid jobs as survey supervisors when not engaged in monitoring), similar levels of cheating by all monitors is unlikely.

entire agricultural season and tapered off before the tree survival monitoring one year after training, consistent with the need for ongoing investments on the part of the farmer.

A note on the timing of farmers' decisions and information. Because the program offered rewards and measured outcomes for one year, farmers' take-up and follow-through decisions are based on their perceptions about costs and benefits during the first year only. The rationale for the reward design is that the costs associated with planting and caring for the trees are highest during the first year when the trees are vulnerable and require attention in the form of watering, weeding and protection from pests. After they survive the first dry season, costs decrease substantially. The follow-through decision we observe is more accurately described as the cumulative outcome from numerous follow-through decisions made over the course of the year after take-up. New information may reveal itself starting immediately after the take-up decision is made, or at different points in time, as family members fall ill, crops fail, or input and output prices change.[22] When new information arrives that affects the opportunity cost of caring for the trees, farmers may reoptimize on the number of trees they continue to cultivate (if any). Note that our empirical model imposes a simplified version of the timing, where we assume there are only two decision periods (take-up and follow-through) as opposed to many. This simplified timing assumption corresponds well to the empirical setting if the bulk of the information arrives shortly after take-up or with a series of shocks that are highly correlated. On the benefit side, we expect to see little change in information within the first year since the private benefits take considerably longer than the costs to materialize. Of course, farmers may still face uncertainty about the costs and benefits of keeping trees alive, even after follow-through is measured.

[22]The take-up decision is made at the beginning of the planting season, as shown in Appendix figure A.5.1. This is the natural timing of take-up decisions for other crops and technologies. Therefore, our design allows for an amount of time between take-up and follow-through that is similar to many other agricultural technologies.

4 Summary statistics and reduced form results

Appendix table A.5.1 shows baseline summary statistics by treatment and treatment balance. Around 70 percent of participants are heads of household and 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. 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 Dunavant Cotton for an average of over 4 years and over 40 percent interact regularly with their lead farmer. Almost 70 percent of respondents report familiarity with the technology but only around 10 percent had adopted prior to the program, likely due in part to the absence of a market for *Faidherbia albida* seeds or seedlings.

We test for balance in the randomization outcomes by correlating observable characteristics with treatment levels and assignment. Appendix table A.5.1 tests balance for the take-up subsidy, threshold performance reward, and surprise reward treatment. Larger households with more non-agricultural assets are more likely to receive lower take-up subsidies on average. Older respondents with larger households and better self-reported soil fertility are marginally more likely to be assigned to the surprise reward treatment. The table consists of 51 separate regressions. Five significant coefficients is therefore consistent with significance threshold of 10 percent.

We also examine whether non-random attrition at any stage of data collection affects internal validity (Appendix table A.5.2).[23] The baseline survey covered over 98 percent of trained farmers, while the end line included over 95 percent of baseline respondents. We

[23]Selection into treatment is also a threat to the experiment's internal validity. By design, this is unlikely: group level participation subsidy treatments were revealed only after individuals arrived for training, and individual-level reward treatments were assigned in a one-on-one interaction with study enumerators.

see some evidence that farmers who received lower take-up subsidies were marginally less likely ($p < 0.10$) to participate in the surveys. Otherwise, survey attrition is balanced across treatments. For the tree survival monitoring, over 95 percent of the 1,092 households that took up the program were located.[24]

Finally, spillovers across treatments pose a threat to the experimental design. Because the take-up subsidy treatment was assigned at the group level, spillovers are relatively unlikely. The value of the threshold reward, on the other hand, varied at the individual level. By revealing the reward value privately to each farmer before the take-up decision, we mitigate the potential that take-up is affected by rewards received by others. However relative reward values may still affect performance since farmers can share information after they leave the training. We test for spillovers associated with the take-up subsidy and the threshold reward and observe little evidence that they affected outcomes (these tests and their findings are reported in Appendix A.4.2).

4.1 Reduced form results

We examine the data for three pieces of reduced form evidence. First, we examine how the incentive offered by the threshold reward affects tree survival outcomes and also indicators of farmer investments in the trees. Second, we look for reduced form evidence consistent with the presence of uncertainty. Third, we briefly address alternative explanations including liquidity constraints and behavioral decision-making.

The effect of economic incentives on follow-through and farmer investments. Table 1 displays means and standard deviations for several program outcomes: take-up, follow-through (tree survival ≥ 35), zero surviving trees and the number of trees conditional

[24] Of the farmers eligible for monitoring, we were unable to locate 9 of them and thus assume zero tree survival in the analysis.

on positive survival rates. These statistics are broken down by treatment and show clear patterns in responses to the incentives offered in the experiment.

The reward amount has a positive effect on follow-through, both in the likelihood that farmers reach the 35-tree threshold and in the absolute number of trees. This can be seen first in Panel C of Table 1, which in column 2 shows that the share of farmers that reached the 35-tree threshold increases from 0.13 to 0.32 across reward groups in ascending order, and in column 3 shows a similarly monotonic relationship between the number of trees and the reward amount. Column 4 also shows that the share of farmers with zero surviving trees falls monotonically with the reward. We also look at the linear relationship between these three outcomes and the reward when we compare our data with the structural estimates (see Table 3). For example, the reward amount (in '000 ZMK) has a marginal effect of 0.044 surviving trees that is significant at the 1 percent level (column 7 of Panel A, Table 3).[25]

Consistent with the follow-through results, we find evidence that farmers' investment choices are responsive to the reward (Appendix table A.5.3). Specifically, enumerators recorded signs of weeding, fire breaks, watering and burning during field monitoring visits at the end of the project. All of these activities are costly to the farmer and are likely to affect tree survival, the first three positively and the last negatively. A linear regression of the probability that the enumerator observed weeding, fire breaks and watering on the threshold reward value shows a positive effect on weeding, fire breaks and watering, with p-values of 0.059, 0.129, and 0.041 respectively. The coefficient on field burning, which threatens tree survival, is negative and statistically insignificant.

[25]The relationship between follow-through and the reward is unaffected by selection into the program based on reward amounts. We can show this by comparing the response to the reward across farmers who learned about the reward after choosing to take-up and farmers that knew about the reward before taking-up. The marginal effect of the reward is statistically similar in the two groups (see Appendix table A.4.1).

Table 1: Summary statistics

	(1) Take-up	(2) 35-tree threshold	(3) # trees | # trees> 0	(4) Zero trees
		Panel A: full sample		
mean	0.83	0.25	27.42	0.36
sd	0.38	0.44	14.31	0.48
		Panel B: by take up subsidy treatment		
A = 0				
mean	0.71	0.26	27.60	0.37
sd	0.46	0.44	14.31	0.48
A = 4000				
mean	0.76	0.29	28.86	0.36
sd	0.43	0.45	13.67	0.48
A = 8000				
mean	0.86	0.27	29.30	0.38
sd	0.35	0.44	14.19	0.49
A = 12000				
mean	0.97	0.22	24.93	0.33
sd	0.17	0.41	14.52	0.47
		Panel C: by reward treatment		
R = 0				
mean	0.90	0.13	22.00	0.49
sd	0.31	0.34	14.70	0.50
R = (0,70000]				
mean	0.90	0.21	25.45	0.40
sd	0.30	0.41	14.62	0.49
R = (70000,150000]				
mean	0.93	0.32	29.53	0.30
sd	0.25	0.47	13.67	0.46

Notes: Means and standard deviations of take-up (column 1) and follow-through (columns 2-4) outcomes, by experimental treatment. Column 1 includes all farmers (N=1314). Columns 2-4 are conditional on take-up (N=1092). Column 2 reports the number of farmers who reached the performance reward threshold.

Reduced form evidence of uncertainty. We use the means and standard deviations presented in Table 1 to provide evidence for the presence of uncertainty, consistent with our conceptual model. Regression-based results are shown, for ease of comparison with simulations from the structural model, in Table 3. First, notice that take-up rates are increasing across values of the take-up subsidy. Take-up rates are high, on average, even in the zero subsidy condition, where over 70 percent of farmers take-up. This could be due to high known payoffs from follow-through, on average, or to high expected values driven by option value (see Proposition 4).

Second, we observe that follow-through rates vary considerably within treatment and are low, on average, with only 25 percent of farmers reaching the 35-tree threshold (column 2). This holds even in the zero subsidy condition, ruling out that farmers were certain about high payoffs associated with cultivating a large number of trees at the time of take-up. Low follow-through conditional on take-up is consistent with Proposition 2.[26]

Third, a large number of farmers abandon the technology altogether (have a survival of zero trees), even conditional on taking up with zero subsidy (37 percent, column 4). This rules out that farmers were certain about positive payoffs from a small number of trees at the time of take-up, as in Corollary 2.1 of our conceptual model.

Finally, we see no reduced form effect of the subsidy treatment on the likelihood of reaching the 35-tree threshold or of abandoning the technology (zero trees). We implement a two-sample t-test for equal means between the highest and lowest subsidy condition. For continuous tree survival, the probability of reaching the threshold (≥ 35 trees) and zero trees, the p-values are 0.63, 0.25 and 0.32, respectively. The linear regression test of the effect of the take-up subsidy on tree survival outcomes (shown in Table 3) is also statistically insignificant. This is consistent with Proposition 3, which states that the selection effect of subsidies will

[26] Behavioral explanations such as over-optimism or procrastination might also be consistent with high take-up and low-follow through, even at positive take-up prices. We discuss behavioral explanations consistent with the reduced form results, as well as the interpretation of the type of new information, in Section 7.

be diminished by high levels of uncertainty in the net benefits of follow-through.

We also examine whether outcomes can be explained by observables. Appendix table A.5.4 shows that, overall, observables explains relatively little of the variation in outcomes: the R-squared from a regression of outcomes on observables is 0.0296, 0.0297 and 0.0314 for take-up, reaching the 35-tree threshold and tree survival, respectively. Adding the treatment variables improves the explanatory power substantially (even numbered columns). The low explanatory power of observables further motivates our use of a structural model to estimate the heterogeneity across farmers at both take-up and follow-through.

Alternative mechanisms. In Appendix 4, we investigate potential alternative mechanisms underlying the reduced form evidence. First, we test whether liquidity constraints had an effect on take-up or self-selection. Second, we investigate psychological channels that may affect both the decision to take-up and to follow-through with the technology. We find little support for the empirical relevance of either explanation.

5 Model, identification and estimation

The reduced form results in Section 4 provide evidence that is consistent with uncertainty in the opportunity costs of follow-through. However, they do not rule out that, in addition to uncertainty, other sources of heterogeneity in costs may explain the lack of screening effect of the take-up cost. For instance, a zero or even negative correlation between follow-through rates and the take-up cost could emerge if there is a negative correlation between the privately optimal number of trees and the total profit farmers derive from them.[27] In addition, the

[27] A correlation (positive or negative) between the optimal scale and the level of profit can emerge from the joint distribution of the primitive parameters that govern a profit function (e.g. marginal costs, fixed costs, marginal benefits, etc.). For instance, Suri (2011) finds that low adoption rates of hybrid maize among farmers who seem to have high returns from adoption can be traced to a positive correlation between fixed costs and marginal benefits from adoption using a random coefficients model.

reduced form results do not offer any insight into the magnitude of the uncertainty that farmers face. To address these remaining questions, we adapt our simple theoretical model described in Section 2 to our empirical setting and explicitly estimate the distribution of random parameters governing a quasi-profit function (a “reduced form” profit function of sorts).

5.1 Farmer net benefits

General profit function. We begin with a general characterization of a farmer profits at time $t = 1$ as a function of the number of trees she decides to plant and care for:

$$\Pi(N) = \left[\sum_{t=7}^{\bar{T}} \frac{1}{(1+r)^t} \left(\alpha_0 N - \alpha_1 N^2\right)\right] - \gamma_0 N - \gamma_1 N^2 - \gamma_2 \times \mathbf{1}(N > 0) \tag{2}$$

where N is the number of trees, the term in brackets corresponds to the discounted flow of benefits, and the remaining terms represent variable and fixed cost. Equation (2) describes a convex function in the number of trees cultivated provided that all parameters are positive and $\tau\alpha_0 - \gamma_0 \geq 0$, where $\tau = \sum_{t=7}^{\bar{T}} \frac{1}{(1+r)^t}$.

The solution to the profit maximization problem defined by (2) is:

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

where $\tau = \sum_{t=7}^{\bar{T}} \frac{1}{(1+r)^t}$. The existence of interior and corner solutions to this function is consistent with two empirical observations: many farmers choose to cultivate zero trees and a number of them find it optimal to cultivate between zero and 50 trees (the number of seedlings they receive) in the absence of an external incentive.[28]

[28] See Appendix figure A.3.1

The differences in tree choices across farmers could emerge from heterogeneity in some or all parameters in (2). Our experimental variation, however, does not allow us to separately identify heterogeneity along all of these dimensions. We therefore turn to a quasi-profit function that uses our experimental variation to characterize farmer heterogeneity along two important dimensions of the farmer's profit maximization problem: the interior solution and the profit level evaluated at the optimal number of trees.

Quasi-profit function and farmer's decision at $t = 1$. The same interior and corner solutions conditions delivered by (2) are generated by the following quasi-profit function indexed by two random parameters, T_i and F_i:

$$\Pi(N) = N - \frac{1}{2T_i}N^2 - F_i \times \mathbf{1}(N > 0) + \mathbf{1}(N \geq \bar{N})R_i \tag{3}$$

where $T_i = \frac{\tau\alpha_0 - \gamma_0}{2(\tau\alpha_1 + \gamma_1)}$, $F_i = \gamma_2 + \left(\frac{(\tau\alpha_0 - \gamma_0) - (\tau\alpha_0 - \gamma_0)^2}{4(\tau\alpha_1 + \gamma_1)}\right)$, and $R_i = 0$. The quasi-profit function in (3) allows for heterogeneity across farmers in two "reduced form" parameters (in the structural sense): T_i, the interior solution, and F_i, which is a scaling parameter that ensures maximum profits in the quasi-profit function coincide with profits in the generic quadratic function. A free correlation structure is key to their interpretation as reduced form parameters, since they are a function of several common structural parameters. A negative or zero correlation between them could generate the type of selection patterns we observe in the data: zero correlation between the take-up cost and reaching the tree survival threshold.

The advantage of the quasi-profit function (3) over (2) is that the joint distribution of its two random parameters is identified out of the variation induced by our experiment: the last term in (3) corresponds to the exogenous threshold reward, which we vary randomly across farmers. The reduced form nature of (3) means that we do not need to specify which structural parameters in (2) are driving the variation in choices. And yet, since (2) and (3) share the same value at the optimal solution, we can still use (3) to evaluate welfare under the more general profit function (2).

Tree survival as a farmer decision. Throughout our estimation, we assume that tree survival is deterministic conditional on farmers' costly effort. As we explain in greater detail in Appendix 3, this assumption is also consistent with a model where survival is probabilistic and the probability of survival is a convex continuous function of effort, e, up to $\tilde{e}$, where it attains one. Farmers would respond to such probability profile by investing the minimum effort that guarantees survival, $\tilde{e}$, in all trees they choose to plant.[29] Empirically, the small bunching of tree survival at 35 (the reward threshold) we observe in the data is consistent with this assumption (see Appendix 3).

5.2 Dynamics and take-up decision

As in the conceptual model, we assume the farmer makes adoption-related decisions in two periods: $t = 0, 1$. The random parameter F_i, which largely determines the magnitude of optimized profits, is divided into two additive components: F_{0i} and F_{1i}, where F_{0i} is known at all periods and F_{1i} is known at $t = 1$ but not at $t = 0$.[30] In addition, we assume that T_i is known to the farmer at all times. This amounts to assuming that there is uncertainty about the net returns to tree cultivation, but not about the optimal scale of the technology. The advantage of this particular structure of information is that it allows us to nest a model without uncertainty (i.e., $\text{Var}(F_{1i}) = 0$) that could also deliver no screening effects (or even negative screening effects) within our more general model.

At $t = 0$, the farmer decides whether or not to pay to take-up the technology. At this point in time, the farmer has partial information about her net benefits from the contract.

[29] Except, perhaps, on one of them, as is explained in Appendix 3.

[30] See the last paragraph of Section 3.2 for a discussion on our two-period assumption.

Assuming the farmer knows the distribution of F_{1i} conditional on F_{0i} and T_i at $t = 0$, the farmer chooses to take-up if

$$\delta \mathbb{E}_{F_{1i}|F_{0i},T_i} \left[\max_N \Pi(N|T_i, F_{0i}, F_{1i}, T_i, R_i) \right] - c + A_i \geq 0 \tag{4}$$

where c is the cost of the seedlings, A_i is the randomly determined subsidy, and δ is the discount factor, assumed equal to 0.6.[31] Note that this representation of the NPV of farmer's profits from trees maintains the risk neutrality assumption from Section 2.[32]

5.3 Identification and estimation

Identification of the structural model consists of uniquely identifying the joint distribution of unobservables T_i, F_{0i} and F_{1i}. In addition to the above described assumptions on the timing of information, we maintain assumptions (i) and (ii) on the components of F_i from Section 2, and add the following assumptions

(iii) No common shocks: $F_{1i} \perp F_{1j}\ \forall i \neq j$

(iv) Normality: $F_{0i} \sim n\left(\mu_F, \sigma^2_{F_0}\right)$, $F_{1i} \sim n\left(0, \sigma^2_{F_1}\right)$

(v) Joint normality: $(F_i, \ln T_i) \sim n(\mu, \Sigma)$

Below we explain the role that each of these assumptions plays for identification. In what follows, we denote the randomized values of A_i and R_i as a_i and r_i to emphasize their role as known (to the farmer and researcher) and exogenous.

[31] Like Stange (2012), we note that in the context of stochastic dynamic structural models the discount factor is not separately identified from the scale parameter of future period shocks. We used survey data on time preferences to inform our choice of 0.6, which is in line with observed interest rates in our setting and elicited individual discount rates in other rural developing country settings (Conning and Udry 2007; Cardenas and Carpenter 2008).

[32] As discussed in Section 2, this assumption is innocuous to the extent that the changes in income produced by our program are small relative to total income. The highest reward from our program is roughly 3.5 percent of average annual income.

With no assumptions other than profit-maximizing behavior on behalf of the farmer and a quadratic profit function that allows for corner solutions, the joint distribution of F_i and T_i can be non-parametrically identified in the subset of the support such that $\bar{N} < T_i < 50$. To see this, consider the follow-through decision of the subset of the sample for which

$$\lim_{a \to \mathcal{A}_1} \Pr\left(\mathbb{E}\left[\max_N \Pi(N|T_i, F_{0i}, F_{1i}, T_i, r_i)\middle| F_{0i}, T_i\right] \geq c - a_i\right) = 1,$$

such that there is no selection on take-up. Within this subset of the sample, we can use the variation in r_i to identify the joint distribution of (F_i, T_i). For this group, the probability of cultivating $N^* = n > \bar{N}$ trees when $R = r_i$ can be written as

$$\Pr(N^* = n; R = r_i) = \Pr\left(F_i < r_i + \frac{1}{2}n\middle| T_i = n\right)\Pr\left(T_i = n\right) \tag{5}$$

Because the left hand side of (5) is empirically observable, increments in r_i, holding n constant, trace out the conditional distribution of F_i given T_i. The same expression can then be used to recover the marginal distribution of T_i by varying n and dividing by the conditional distribution of F_i. Since non-parametric identification of the joint distribution of F_i and T_i occurs only in the subset of the support such that $\bar{N} < T_i < 50$, additional parametric assumptions are required to fully characterize these distributions. We therefore adopt assumption (v) for the estimation.

We use farmers' take-up decisions in combination with assumptions (i)-(iv) in order to separately identify the distributions of F_{0i} and F_{1i}, once the joint distribution of F_i and T_i has been identified. Under these assumptions, the decision to take-up in response to r_i and a_i provides independent identification of the distribution of the known component of F_i, F_{0i}. More formally, identification of the distribution of F_{0i} is obtained from the decision to take-up, which is characterized by the inequality in (4). The left side of (4) is a known function of the random variable F_{0i}. Note that parameters $\mu_F, \sigma_F^2,\ \mu_T,\ \sigma_T^2,$ and $\rho_{T,F}$ can be treated as known since they are identified from tree survival as described above. Denote this function $h(F_{0i}; r_i)$, so we can rewrite (4) as

$$h(F_{0i}; r_i) \geq c - a_i \tag{6}$$

The right side of (6) can take one of four known values, as $a_i \in \{0, 4000, 8000, 12000\}$. The left hand side of (6) is known up to F_{0i} and varies across individuals in response to the known cost determinant, r_i. Provided that $h(F_{0i}; r_i)$ is invertible,[33] we can identify the distribution of F_{0i}, from the random variation in a_i and r_i:

$$\Pr\left(F_{0i} \leq h^{-1}(c - a_i, r_i) | \mu_F, \sigma_F^2, \mu_T, \sigma_T^2, \rho_{T,F}\right) = \Pr(TakeUp_i | a_i, r_i) \tag{7}$$

Common shocks and mean shift model. Assumption (iii) plays an important role for identification, as it implies that the variance of F_i across farmers is the sum of the variances of its two components: $\sigma_{F_0}^2 + \sigma_{F_1}^2$.[34] The variance of shocks is thus partially identified from subtracting the variance estimate of F_{0i}, identified from (7), from the variance of F_i, identified from tree choices. Shocks that are common across farmers do not translate into variance in tree choices, and would lead to an underestimate of $\sigma_{F_{1i}}^2$. In our context, much of the uncertainty farmers face appears to be from idiosyncratic shocks. According to our survey, two-thirds of respondents list health problems as their greatest challenge, almost 50 percent of households report losing cattle or livestock to death or theft during the past year, and 10 percent of households report the death or marriage of a working age member.[35] However, given that farmers are also likely affected by common shocks such as

[33] It can be shown that there exists some $\bar{f}$ s.t. $h(F_{0i}; r_i)$ is strictly monotonically decreasing on $(-\infty, \bar{f})$.

[34] Assumption (iii) is also present in Fafchamps (1993), and is necessary for maximum likelihood estimation.

[35] This is consistent with a literature that documents, in most cases, a disproportionate share of income risk from idiosyncratic factors in rural developing country settings (summarized in Dercon 2002).

rainfall patterns and commodity prices, we estimate a variant of our model that allows for a specific type of common shock: one that is completely unforeseen at the time of take-up and is common across all farmers. This model variant can be estimated by relaxing assumption (ii), i.e. allowing for subjective and objective expectations about the mean of the shock to differ. Thus we refer to this model as the *mean shift model*. The mean shift – or difference between the subjective mean of the shock distribution at take-up and its objective mean – is identified because the random variation in r_i and a_i allows us to identify μ_F in (7) from the take-up decisions independently of the estimate from tree choices. Besides allowing us to incorporate a type of common shock, the mean shift also captures any over-optimism or experimenter demand effects that are common across farmers. These behaviors will share the same structure as the unforseen common shock: the subjective mean will differ from the objective mean of the opportunity cost distribution.

Estimation. We estimate the model using simulated maximum likelihood. The log-likelihood function is over observations of the number of planted trees, $N = 0, ..., 50$, and the participation decision, $DP = 0, 1$. The sample includes the 1,314 farmers who made a take-up decision. Because there are no trees planted whenever the individual chooses not to participate, the support of this bivariate vector is given by the 52 (DP, N) pairs: $(0, 0), (1, 0), (1, 1), (1, 2), ..., (1, 50)$.

$$\begin{aligned} l(\xi; DP, N) \quad &= \\ &\textstyle\sum_{i=1}^{M} \left\{ (1-DP_i)\ln(1-\pi_{P,i}) + DP_i \ln(\pi_{P,i}) + DP_i \sum_{j=0}^{50} \mathbf{1}(N_i = j) \ln \Pr(N = j) \right\} \end{aligned} \tag{8}$$

where $\xi = (\mu_F, \sigma^2_{F_0}, \sigma^2_{F_1}, \mu_T, \sigma^2_T, \rho_{T,F})$.

We use numerical methods to minimize the negative of the simulated log-likelihood. For each likelihood evaluation, we use 1,500 draws of (T_i, F_{0i}, F_{1i}). Within each likelihood evaluation and for each draw of (T_i, F_{0i}, F_{1i}), the expectation on the right hand side of equation (4) is numerically computed using 100 draws of (T_i, F_{1i}) conditional on the draw of F_{0i}.[36] Standard errors for the estimated parameters are obtained as the inverse of the inner

[36]Simulated methods often result in stepwise objective functions which work poorly with gradient-based

product of the simulated scores.[37]

6 Structural estimates and simulation results

In this section, we describe the structural estimates and carry out counterfactual simulations.

6.1 Structural estimates and model fit

Table 2 shows the point estimates for the main parameters described in Section 5.3.[38] Panel A shows the estimates of our baseline model, which assumes that farmers' expectations about F are correct and consistent over time. Panel B shows the results of allowing for an unexpected common shock to all farmers at $t = 1$ (a mean shift). Because point estimates are somewhat hard to interpret (e.g. the μ_T and σ_T parameters do not correspond to the mean and standard deviation of the log-normally distributed parameter T), we convert the estimated parameters

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. A further discussion of the estimation algorithm can be found in Appendix A.2.

[37] See Appendix A.2 for a more detailed description of our standard error calculation.

[38] There are two remaining parameters that are omitted from Table 2 but discussed in Appendix A.2 for the sake of brevity: these are the surprise treatment parameter, α_S, and the monitoring treatment parameter, α_m. Recall that farmers in the surprise reward treatment made a take-up decision before learning their reward. We model this aspect of the design by assuming individuals expect a threshold reward of 0 when their participation decision is made, but incorporate the reward value they draw in their follow-through decision. Because our reduced form results show that individuals in the surprise treatment had higher participation rates than those individuals who drew a reward of zero ZMK, we allowed the surprise reward treatment to have an independent effect on the participation decision. The structural estimates suggest that the boost to participation is equivalent to offering them between 92 and 54 ZMK (in the base and mean shifter models, respectively). Appendix A.2 describes these results in more detail.

In all models, we allow the regular visits to collect data on program implementation that were administered to one-fifth of farmers to independently affect the tree survival decision (but not the participation decision). The estimated parameter is -238.40 (s.e. 73.887) in Panel A and -229.53 (s.e. 74.444) in Panel B. In other words, regular monitoring visits appears to be reducing the fixed costs of tree cultivation, which is consistent with the positive effect of monitoring on tree survival that we find in the reduced form results.

into more easily interpretable outcomes using simulation.[39] The estimated joint distribution of T and F shown in Panel A is such that the mean ex-post privately optimal number of trees is 8.46 (s.d. 14.64), with about 59 percent of farmers choosing to plant no trees.[40]

Table 2: Structural parameter estimates

Parameters in T					Parameters in F			
μ_T	σ_T	ϱ	μ_F	σ_{F0}	σ_{F1}	α_s	α_m	μ_{Fs}
			Panel A. No Mean Shift					
3.539	1.401	0.818	107.58	307.87	211.42	-91.79	-238.40	-
(0.057)	(0.066)	(0.066)	(11.822)	(93.278)	(49.953)	(16.222)	(73.887)	-
			Panel B. Allowing for Mean Shift					
3.579	1.392	0.835	74.48	290.06	193.05	-54.42	-229.53	53.29
(0.071)	(0.075)	(0.073)	(15.47)	(84.622)	(45.427)	(20.47)	(74.444)	(26.761)

Notes: Parameters fitted by simulated maximum likelihood using 1500 draws of the random vector (F_{0i}, F_{1i}, T_i), with smoothing (lambda is 0.5) and tolerance (1e-15). The baseline model (Panel A) restricts the mean of F_i to be the same in both time periods. The mean shift model (Panel B) allows the mean of F_i to differ between the two periods, and the parameter F_shift to capture this difference. The log-likelihood value for the baseline model is 11142.064, while it is 11138.996 for the mean shift model.

This joint distribution also implies that the average ex-post private profit from the optimal number of trees is 108.39 thousand ZMK. However, ex-post private profits vary widely across farmers: the s.d. is 185.47. Importantly, the model estimates that about 39 percent of the variance in ex-post profits is due to new information that emerges after the take-up decision is made. To put this magnitude into perspective, among farmers whose expected number of trees is above the reward threshold at the time of take-up, around 15 percent prefer a number below the threshold after new information is revealed.

The variance of shocks is partially identified out of the difference in the variances of

[39] The point estimates in Table 2 can be used to simulate farmer's draws of F and T. These draws are then used to compute optimal tree cultivation decisions that account for interior and corner solutions, which are then be plugged back into the profit function to compute maximized profit. The statistics presented here correspond to means and variances from 10,000 simulated draws.

[40] These statistics assume that farmers can plant a maximum of 50 trees. Although we allow for the distribution of T to be unbounded, we present statistics of the bounded distribution because we fit the econometric model using only this range of outcomes. According to our estimates, about 6 percent of the farmers would choose a private optimum of 50 or more trees.

expected profits at take-up ($t = 0$) and ex-post profits at the time the follow-through decision is made ($t = 1$) and partially identified out of its non-linear effect on the mean level of the expected profits at take-up.[41] The presence of common shocks will generate a tug-of-war between these two sources of identification: the expected value of the profits will pull the variance of shocks parameter, σ_{F1}, up while the ex-post variance in profits will not reflect the variance of common shocks and thus will pull σ_{F1} down. Our mean shift model helps us explore the importance of the bias in σ_{F1} generated by the presence of common shocks, by allowing the mean level of profits to be different at $t = 0$ and $t = 1$.

The corresponding results from the mean shift model are presented in Panel B of Table 2. The estimated difference in means between the two periods, the mean shift, is given by parameter μ_{F_S}. A non-zero value for this estimated parameter has a couple of plausible interpretations. First, it can represent a single common shock whose possibility was unknown at the time of take-up and affected all farmers equally. Second, it can pick up a common update in the value for the technology that occurred after the take-up decision was made. The latter interpretation is a useful test for the presence of experimenter demand effects on take-up: the perceived obligation of potential adopters to take-up in the presence of the experimenter. Our estimate of the value for μ_{F_S} is positive but small and not significantly different from zero at standard confidence levels. The positive value is consistent with either the presence of a common shock of the specific type described above or an experimenter demand effect. Note, however, that in either case, it is small compared to the standard deviation of the shocks, σ_{F_1}. And, importantly, allowing for this effect induces a small change in the variance of the shocks: σ_{F1}, falls moderately from 211.42 to 193.05, which is consistent with a positive bias in our baseline model stemming from the presence of common shocks.[42] This suggests that uncertainty in the form of idiosyncratic shocks is important and

[41] Recall that a higher variance in shocks results in a larger option value for the contract.

[42] Unfortunately, we cannot calculate the share of the variance attributed to common shocks using model estimates since we are not explicitly modeling random common shocks with a known distribution at the time

that experimenter demand effects are not driving our results.

We now turn to the interpretation of the parameters that govern the distribution of known (at the time of take-up) sources of heterogeneity across farmers. We estimate a high positive correlation between F and $\ln T$, $\rho_{T,F} = 0.81$ ($\rho_{T,F} = 0.83$ in the mean shift model).[43] Because F enters negatively in the profit function, this translates into a negative correlation between the privately optimal number of trees and the level of profits. This negative correlation generates low, but strictly positive, follow-through rates (as in low numbers of trees planted) among farmers whose expected value from the contract is high and thus take-up under a high price. In other words, the static heterogeneity identified by the model contributes to undermining the effect of high prices at take-up on positive (high follow-through) self-selection of farmers.

Because F and T are reduced form parameters (see Section 5.1), the positive correlation between them could stem from two sources: a positive correlation between farmers' fixed costs and farmers' interior solution to the maximization problem (a component of the reduced form random parameter, F, is the fixed cost of the generic quadratic profit function), or a high mean and high variance in the linear term of net marginal returns (the term $\tau\alpha_0 - \gamma_0$ in the reduced form expression for T and F), which enters linearly in T and non-linearly in F, and in high ranges may generate a negative relationship between them.[44] Our model predicts well some simple observations in the data. For example, our baseline model predicts that 1,104 farmers (1,112 under the mean shift model), out of a total of 1,314, will take-up; our data show 1,092 participants. Our model also predicts that 173 out of the 963 farmers that faced a strictly positive take-up cost (i.e. a subsidy less than 12,000 ZMK) will choose to cultivate zero trees (180 in the mean shift model), while the observed number of farmers in

of take-up.

[43] F_1 is assumed to be orthogonal to F_0 and T. Thus, the correlation between F and T stems solely from the correlation between F_0 (the known component of F) and T.

[44] We verified this is the case for plausible distributions of the deep structural parameters via simulation.

this category is 112. That is, the dynamics in our model replicate an observation that seems at odds with rationality in a static framework: some farmers who purchase the trees choose not to cultivate them. In this sense, our result parallels the result by Fafchamps (1993) in that individuals make costly choices to maximize future flexibility.

Next, we further examine the model fit by comparing the reduced form treatment effects using simulated outcomes from the estimated model and the observed data. Most of the magnitudes and signs between the treatments and outcomes are well matched by our model estimates. Panel A of Table 3 shows results with the observed data, while Panels B and C show the corresponding simulations using the estimates from Table 2. Columns 1-4 estimate the effect of a thousand ZMK increase in the subsidy on take-up and follow-through outcomes. Columns 5-8 repeat the regressions with the reward (in '000 ZMK) on the righthand side. The effect of the subsidy and the reward on take-up (columns 1 and 5) and whether the farmer reached the 35-tree threshold (columns 2 and 4, conditional on take-up) are similar across the observed and simulated data. The effect of the reward on zero tree survival is also very similar in the observed and simulated data.

There are, however, some discrepancies between what our model predicts and the observed data. Columns 3 and 7 show the effect of the take-up subsidy and reward on the number of surviving trees, excluding zeros (which are shown in columns 4 and 8). Interestingly, the sign of the coefficient on the subsidy is different between the observed and simulated data, though standard errors are relatively large. The effect of the reward on the number of trees (column 7) is larger in the simulated data, consistent with the effects on reaching the 35-tree threshold. Finally, we see evidence that the take-up subsidy selects for farmers more likely to keep zero trees alive (column 4) in the simulated but not in the observed data, indicating some selection effect on abandoning the technology altogether in the simulated data only. Using simulations, we explore whether some of the discrepancies between our estimated model and the data stem from the assumption of deterministic survival of the

Table 3: Comparison of structural and reduced form estimates

	(1) Take-up	(2) Follow-through	(3) # trees | # trees>0	(4) 1.(zero trees)		(5) Take-up	(6) Follow-through	(7) # trees | # trees>0	(8) 1.(zero trees)
					Panel A. Observed Data				
Take-up subsidy	0.022***	-0.004	-0.229	-0.003	Reward	0.001*	0.001***	0.044***	-0.001***
	(0.005)	(0.004)	(0.200)	(0.005)		(0.000)	(0.000)	(0.013)	(0.000)
Observations	1,314	1,092	701	1,092		624	1,092	701	1,092
R-squared	0.071	0.002	0.005	0.001		0.006	0.018	0.022	0.019
					Panel B. No Mean Shift				
Take-up subsidy	0.020***	-0.003	0.044	0.010***	Reward	0.001*	0.003***	0.092***	-0.001***
	(0.002)	(0.003)	(0.130)	(0.003)		(0.000)	(0.000)	(0.012)	(0.000)
Observations	1,314	1,126	631	1,126		624	1,126	631	1,126
R-squared	0.064	0.001	0.000	0.008		0.005	0.105	0.083	0.011
					Panel C. Allowing for Mean Shift				
Take-up subsidy	0.020***	-0.003	-0.002	0.008**	Reward	0.001*	0.003***	0.094***	-0.001***
	(0.002)	(0.003)	(0.124)	(0.003)		(0.000)	(0.000)	(0.012)	(0.000)
Observations	1,314	1,120	605	1,120		624	1,120	605	1,120
R-squared	0.062	0.001	0.000	0.006		0.006	0.107	0.089	0.013

Notes: This table shows coefficients from regressions of each of four indicator variables (take-up, follow-through (threshold), tree survival larger than zero, and no tree survival) on each of our randomized treatments (take-up subsidy and threshold reward). Panel A shows these regression outcomes for the true data. Panel B and C show the fit of the structural model by simulating all four outcomes using the model estimates and examining the how much the linear relationships between outcomes and treatments resemble those in Panel A. Panel B uses baseline model estimates (Panel A of Table 2), while Panel C uses estimates from the mean shift model (Panel B of Table 2). Column 5 is estimated for farmers who learned about the reward before making their take-up decision only.

trees conditional on effort by introducing a stochastic component to tree survival outcomes, taking parameter estimates as given. We find little to no improvement when stochasticity is introduced into the tree survival outcomes (see Appendix A.3).

6.2 The effect of uncertainty on farmer profit and program outcomes

We use estimates from the structural model to perform counterfactual simulations of farmer profits and program outcomes (take-up and tree survival) under different levels of uncertainty. For these analyses, we use the results from the model with the mean shift parameter, such that both the mean and the standard deviation of F_i may vary over time (Panel B in Table 2). In the simulated results below, we set the mean shift for $t = 1$, μ_{F_s}, equal to zero, so that we can equate the expected benefits with the true average discounted benefits from the program.[45] Results from using our baseline model estimates are qualitatively similar.

Farmer profits

We begin by examining the effect of uncertainty on the average per-farmer expected private profit (right hand side of inequality 1) implied by the empirical model. In order to do so, we simulate the value of the expected profit for each farmer at different values of σ_{F1}. The relationship between the mean expected profit and σ_{F1} corresponds to the solid black curves in Figure 2. Panel A shows the relationship for a reward of zero and Panel B for the largest reward offered (150 thousand ZMK). Both are shown at a full subsidy, so that take-up is 100 percent (i.e. there is no selection on the take-up price).

[45] This treatment of the mean shift parameter is consistent with the common-shock interpretation of this parameter.

Expected farmer profits are increasing in uncertainty. This result is analogous to the theoretical result discussed in Proposition 4: the option value from the contract increases with uncertainty and thus drives a positive relationship between the expected profit and uncertainty. The option value, as defined in Appendix 1, is shown by the dashed lines in Figure 2 for different values of the reward. The option value is always non-negative, and is also the only component of the expected private profit that varies with uncertainty. Recall that this result stems from the asymmetric response of the expected profit to positive and negative shocks: if the realization of the random component of the profit drives it below zero, the farmer will respond by not cultivating any trees at all (effectively bounding the profit realizations at zero). This optimizing behavior turns the high variance of the shocks into an asset of the contract, which in turn results in higher take-up.

The positive relationship between expected private profit and uncertainty has implications for take-up decisions: under higher uncertainty, more farmers are ex ante attracted to the contract, even though its ex post expected value is unchanged. A high enough level of uncertainty may result in an expected profit that exceeds even an unsubsidized take-up cost. Hence, the ability of the take-up cost to "tease out" those who will be more likely to reach the tree survival threshold decreases with uncertainty. This can be observed more directly when looking at take-up and follow-through outcomes as a function of uncertainty, which we turn to next.

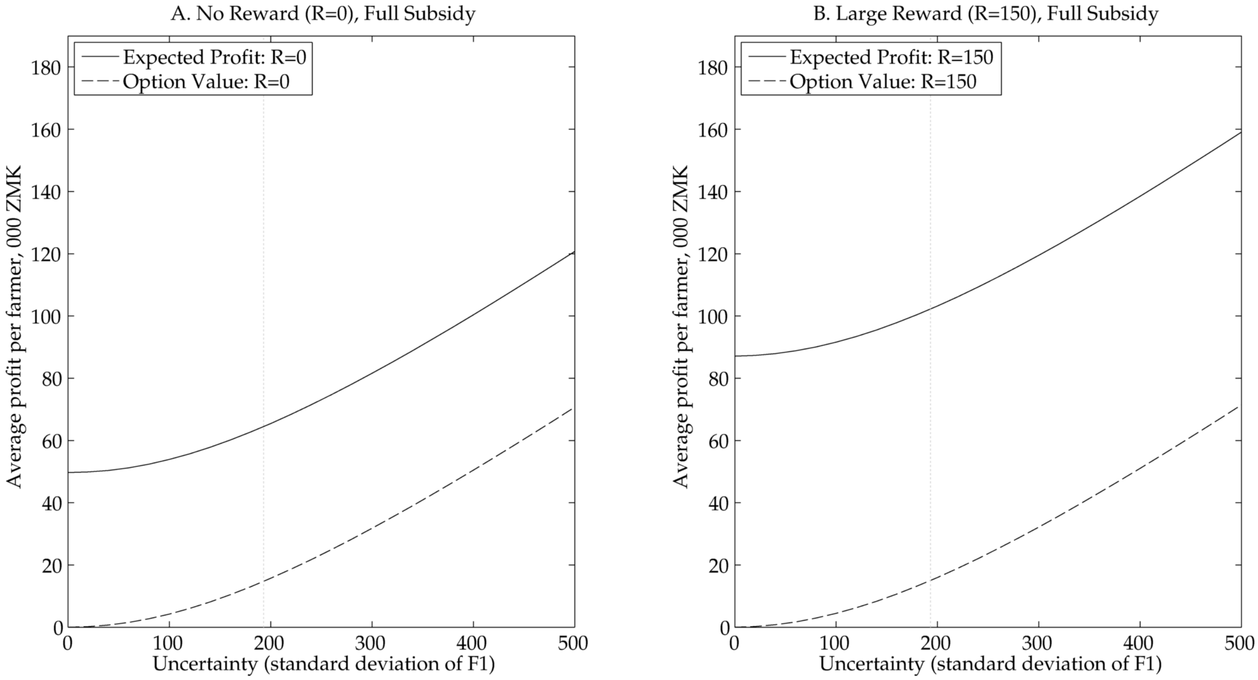


Figure 2: Farmer expected profit as a function of uncertainty

Notes: This figure shows a counterfactual simulation of farmers' mean expected profit as we vary the level of uncertainty (the standard deviation of F_1). For the simulations, we use the estimated parameters from Panel B of 2, except for σ_{F_1}, which we vary along the horizontal axis. The mean per-farmer profit is shown in a solid black line for low (Panel A, $R = 0$) and high (Panel B, $R = 150$) reward values. Profits, the subsidy (A) and the reward (R) are expressed in thousand ZMK. The dashed lines show the mean option value for the two different reward levels. We define the option value as the value of re-optimizing the number of trees to cultivate after F_1 is realized relative to the value of a static choice.

Program outcomes

Figure 3 plots average take-up at low and high subsidies (dashed and solid lines) and low (Panel A) and high (Panel B) rewards as a function of uncertainty. Panels C and D show the share of individuals who reach the 35-tree threshold conditional on take-up for the same combinations of subsidies and rewards. Figure 4 shows the effects of uncertainty on the average number of trees for different values of the subsidy (A) and reward (R). Panels A and B show the average tree survival, unconditional on take-up (non-participants have zero surviving trees). Panels C and D show average tree survival conditional on take up. Because take-up is 100 percent with a full subsidy (A=12), the solid lines in Panels C and D coincide with the solid lines in the Panels A and B, respectively.

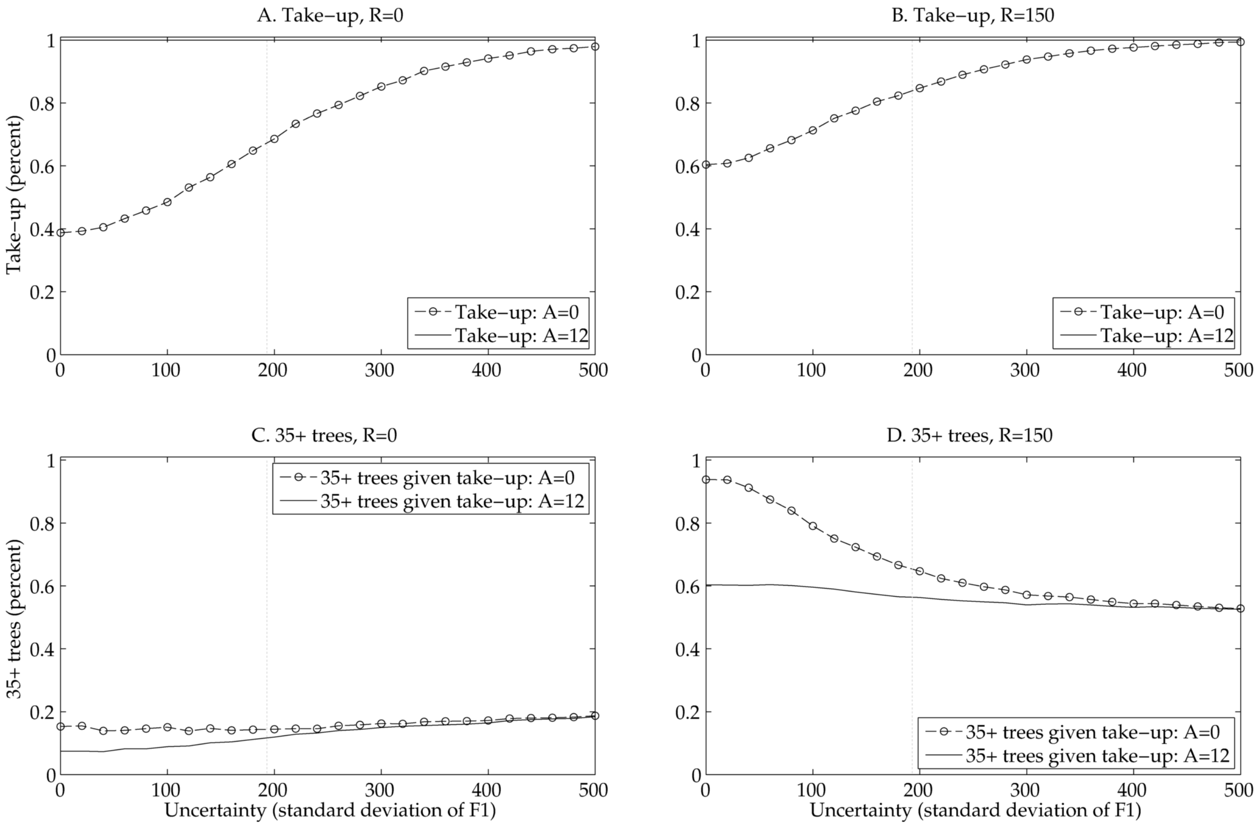


Figure 3: Take-up and threshold outcomes as a function of uncertainty

Notes: This figure shows a counterfactual simulation of farmers' average take-up and 35-tree reward threshold outcome as we vary the level of uncertainty (the standard deviation of F_1). For the simulations, we use the estimated parameters from Panel B of Table 2, except for σ_{F_1}, which we vary along the horizontal axis. Take-up and threshold (trees $\geq$ 35) outcomes are shown for different combinations of the reward value and take-up subsidy, both of which are shown in thousand ZMK.

When uncertainty is low, a higher take-up price increases tree survival conditional on take-up. This result can be seen from the difference between the two lines in Panels C and D, and is what we refer to as the selection effect of price. The existence of a selection effect for low levels of uncertainty, is analogous to Proposition 1 in our conceptual model (Section 2). For the level of uncertainty identified from our data ($\sigma_{F1} = 195$), the gain in tree survival from charging more at take-up is modest – less than 5 trees – and it continues falling as uncertainty increases. The reduction in the selection effect as uncertainty increases is analogous to Proposition 3 in our conceptual model.

When the price for take-up is high, uncertainty lowers follow-through conditional on take-up. This result is shown by the downward slope of the dashed line with circles in Panels C and D in Figure 4. This reduction in tree survival is driven by the selection effect of the take-up price decreasing as uncertainty increases: as shown by Panels A and B in figure 3, take-up converges to 100 percent as uncertainty increases, making the pool of takers ever more similar to the overall pool. This is broadly consistent with Proposition 2 from our conceptual model, although the result in this richer model is qualified by an additional competing effect: the variance of shocks can have a positive or negative effect on the number of farmers that choose zero trees depending on whether the sign of this corner solution threshold is to the right or to the left of the mean of F. We call this effect *the corner solution effect.* It is positive whenever the threshold reward is 150 ZMK, and negative whenever the threshold reward is zero. Thus, when the reward is zero, we see that the unconditional average number of trees increases slightly with uncertainty (both lines in Panel A and the solid line in Panel C increase with σ_{F_1}).[46] Thus, in this richer model, the effect of uncertainty on follow-through conditional on take-up carries both effects: the selection effect, which lowers follow-through; and the corner solution effect, which has an ambiguous effect on follow-through. In our context, the selection effect dominates the corner solution effect, and thus follow-through falls with uncertainty conditional on take-up. That average

[46] This effect holds whenever the distribution of shocks is continuous and symmetric around the mean. To see this, denote the standard normal cdf as $\Phi(.)$ and the threshold along the support of F above which the private profit associated with the interior solution is negative as $\tilde{F}$. The probability of choosing to plant zero trees takes the form $1 - \Phi\left(\frac{\tilde{F}-\mu_F}{\sigma_{F_0}+\sigma_{F_1}}\right)$. Note that the derivative of this probability with respect to σ_{F_1} has the same sign as the numerator of the argument of $\Phi(.)$. When $R = 0$, $\tilde{F} = \frac{1}{2}T_i$. Given that μ_F is above 100 and the mode of T_i is 8.9, the numerator tends to be negative. However, when $R = 150$ and $N^* \geq 35$, $\tilde{F} = N^* - \frac{1}{2T_i}N^{*2} + 150$; and, thus, the numerator tends to be positive.

follow-through conditional on take-up falls with uncertainty has implications for technologies whose benefits kick in above a certain "usage" level: under low levels of uncertainty about implementation costs, a high cost of take-up will help select those individuals who are likely to engage in more intensive usage of the technology.

The effect of a high subsidy on take-up may dominate its effects on selection. This result is most clearly seen by comparing unconditional tree survival (Panels A and B of Figure 4) with tree survival conditional on take-up, which excludes the take-up effect (Panels C and D). The boost associated with the selection effect observed at low levels of uncertainty in Panels C and D is more than offset by the take-up effect: many more farmers take-up when subsidies are high (see Panels A and B in Figure 3). The two counteracting effects lead to similar average tree survival across subsidy levels, unconditional on take-up (Panels A and B). Hence, for technologies whose benefits kick in with total follow-through (whether or not follow-through is spread about few or many adopters), subsidies may increase the benefits of adoption. Note, however, that uncertainty also lowers the take-up advantage of high subsidies, as take-up increases with uncertainty for all subsidy levels.

When uncertainty is high, a reward conditional on follow-through is more effective at inducing tree survival than a subsidy. This result can be appreciated when comparing the effect of moving from a lower R to a higher R as compared to moving from a low A to a high A. Even with the optimal (in our setting) combination of a low take-up subsidy and high threshold reward, uncertainty can bring down the number of farmers that reach the 35-tree threshold.

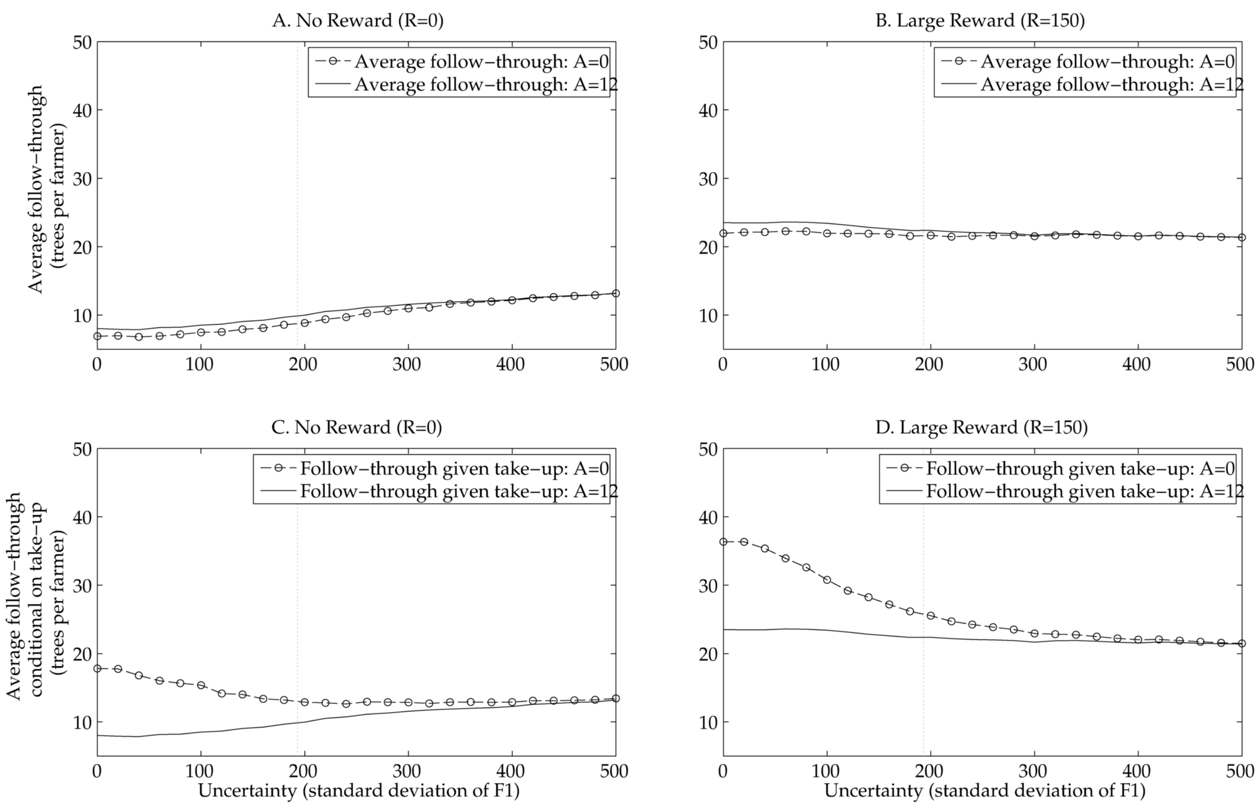


Figure 4: Tree survival as a function of uncertainty

Notes: This figure shows a counterfactual simulation of farmers' mean number of surviving trees as we vary the level of uncertainty (the standard deviation of F_1). For the simulations, we use the estimated parameters from Panel B of Table 2, except for σ_{F_1}, which we vary along the horizontal axis. The mean number of surviving trees is shown for different combinations of the external threshold reward and the subsidy for take-up, both of which are shown in thousand ZMK. The top panels show per-farmer tree survival for all farmers (those who didn't take up have zero trees); the bottom panels show tree survival conditional on take-up.

7 Discussion and interpretation

Our model and results are consistent with several interpretations of intertemporal adoption decisions and the nature of the uncertainty that farmers face. Our preferred interpretation is one of idiosyncratic and common shocks to the opportunity cost of follow-through with the technology, with idiosyncratic shocks playing a relatively more important role. We describe supplementary evidence on alternative interpretations and explanations in this section.

7.1 Learning

Our theoretical and empirical models are similar, from an ex ante perspective, if the information that arrives between $t = 0$ and $t = 1$ is a persistent (learning) or transient shock to opportunity cost, provided that both types of new information are independent across farmers. In addition, to the extent that learning shocks are common across all farmers and unexpected, the mean shift model could account for learning. Interpreted this way, the positive mean shift estimate is consistent with farmers systematically underestimating the costs of follow-through at the time of take-up. The small estimated magnitude of the mean shift parameter would then imply little systematic updating across all farmers. Learning shocks that are independent across farmers and whose distribution is known by the farmers ex ante would show up in the variance of F_1. Thus, we cannot distinguish between persistent and transitory shocks to the opportunity costs of cultivating trees.

The potential for learning about the value of the technology during the first year of tree cultivation is limited given that tree survival benefits are not realized until several years after the end of the study, and planting and caring for trees resembles activities that farmers undertake regularly. Thus, if learning occurs, it is likely related to the opportunity costs of cultivating the trees, rather than the benefits. We use survey data to explore the extent to which baseline knowledge of the technology affects the results. Specifically, we expect that farmers who had more baseline knowledge have less to learn and are therefore less likely to end up with zero surviving trees if they take-up under a positive price.

We construct three baseline knowledge measures: (1) whether the farmer had any *Faidherbia albida* on their land at baseline, (2) whether anyone in the farmer group had *Faidherbia albida* on their land at baseline, and (3) whether the farmer could name at least one risk to tree survival at baseline. Appendix table A.5.5 shows the results from a linear regression of the likelihood of having zero surviving trees on a knowledge variable, an indicator for paying a positive price and their interaction, conditional on take-up. Learning would predict a negative and significant interaction. We observe a negative and statistically insignificant interaction for the first two knowledge variables, and a positive and significant interaction for the third. In other words, we find no clear evidence that learning explains the pattern of take-up under a positive price coupled with zero surviving trees, which can be explained by transient shocks that arrive after take-up.

7.2 Procrastination

An alternative explanation that would violate our identifying assumptions is procrastination or hyperbolic time preferences. Sustained effort choices are frequently associated with time inconsistent behavior, in which the individuals initially takes up, intending to follow-through. But when the time comes to act, costs loom larger (or benefits smaller) than anticipated at the time of initial take-up decision. 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 by constructing two measures of procrastination from the survey data (see Appendix Table A.5.6), which differentiate to some degree between naive and sophisticated 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 subsidy level. For the self-aware (sophisticated) 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 suggest a relatively minor role for procrastination in driving take-up or follow-through outcomes.

8 Conclusion

This paper shows that uncertainty can play an important role in the adoption of technologies that require costly effort over time. We provide a broad framework for adoption decisions that allows for both time-invariant and time-varying heterogeneity as well as multiple dimensions of static heterogeneity across potential adopters. This framework applies to many adoption decisions in agriculture, development, environment, and health research (for example, the adoption and adherence to medical treatments). In our conceptual model, we show that uncertainty in the opportunity cost of adoption can increase take-up rates at the cost of reducing average follow-through rates. Uncertainty at the time of take-up provides an additional explanation to what has been discussed in the technology adoption literature for why charging higher prices may be ineffective at selecting for adopters likely to follow-through at the time of take-up.

Findings from our field experiment show reduced form evidence that is consistent with our conceptual model, in the context of agricultural technology adoption. In Zambia, farmers decide whether to take-up a nitrogen fixing tree under considerable uncertainty about the benefits and costs of following through to keep the trees alive. 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 structural model also helps distinguish between static and time-varying sources of heterogeneity that may explain the absence of screening effects of prices. We find that, in our setup, both static and time-varying heterogeneity reduce the screening effect of the take-up price. Counterfactual simulations indicate that reducing the uncertainty in our setting by half would increase the number of farmers that reach tree survival threshold by 18 percentage points (36 percent).

Like all empirical case studies, our data are specific to our setting. However, the combination of the experimental data with a structural model allows us to simulate adoption outcomes under different levels of uncertainty. Our results are consistent with more than one interpretation of what changes for farmers between the time of the take-up and the follow-through decisions. First, farmers may experience shocks, such as an illness in the household or the arrival of agricultural pests, that affect the opportunity cost of following through with the technology. Second, farmers may acquire additional information about the net costs of tree cultivation after take-up through learning by doing or learning from peers. From an ex ante modeling perspective, our general framework is agnostic about which explanation is correct. Though we cannot distinguish among them using our experimental design, supplementary evidence suggests that idiosyncratic shocks are important in our setting and that learning opportunities, while present, are minimal.

Our study is an example of how experimental variation can be used to identify dynamic structural models. The use of experimental variation in treatments at two different points in time offers an alternative to a panel data structure, since statistically independent samples are exposed to different treatment combinations. To our knowledge, this is the first paper to introduce experimental variation in order to satisfy the exclusion restrictions needed for sequential identification. One caveat of our basic identification strategy is that it relies on shocks being independent across farmers. Therefore, a variant of our model allows for a uniform common shock to farmers, provided that it is completely unanticipated (i.e. the subjective probability assigned by farmers is zero).

From a policy standpoint, uncertainty has the effect of lowering adoption outcomes per dollar of subsidy invested, while increasing the expected private profits to the adopter, because the downside risk of take-up is bounded at zero. To the extent that subsidies rely on public funds, an increase in uncertainty represents an ex ante transfer from the public to the private domain, driven entirely by the adopter's ability to re-optimize follow-through once new information becomes available. While stronger contracts that force adopters to follow-through once they take-up a subsidized technology would address the problem of high take-up coupled with low follow-through, they would do so at a clear cost to the adopter. Future research to explore more innovative solutions to encouraging both take-up and follow-through in the presence of uncertainty offers a promising direction for both environmental and development policies. For example, cheaper monitoring solutions that facilitate rewards for follow-through outcomes can have positive effects on both take-up and follow-through, as shown in our setting.

References

Akinnifesi, F. K., O. Ajayi, Sileshi, P. W. Chirwa, and J. Chianu (2010). Fertiliser trees for sustainable food security in the maize-based production systems of East and Southern Africa: A review. *Agronomy for Sustainable Development 30*(3), 615–629.

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

Ashraf, N., J. Berry, and J. M. Shapiro (2010). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review 100*(5), 2383–2413.

Ashraf, N., B. K. Jack, and E. Kamenica (2013). Information and subsidies: Complements or substitutes? *Journal of Economic Behavior & Organization 88*, 133–139.

Barnes, R. and C. Fagg (2003). Faidherbia albida: Monograph and annotated bibliography.

Beaman, L., D. Karlan, B. Thuysbaert, and C. Udry (2014). Self-selection into credit markets: Evidence from agriculture in mali. *Working Paper*.

Brown, S. (1997). *Estimating biomass and biomass change of tropical forests: a primer*, Volume 134. Food & Agriculture Org.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica 82*(5), 1671–1748.

Cardenas, J. C. and J. Carpenter (2008). Behavioural development economics: Lessons from field labs in the developing world. *Journal of Development Studies 44*(3), 311–338.

Carter, M. R., R. Laajaj, and D. Yang (2014). Subsidies and the persistence of technology adoption: Field experimental evidence from mozambique. *Working Paper*.

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics 125*(1), 1–45.

Cohen, J., P. Dupas, and S. Schaner (2015). Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial. *The American Economic Review 105*(2), 609–645.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in ghana. *The American Economic Review*, 35–69.

Conning, J. and C. Udry (2007). Rural financial markets in developing countries. In *Handbook of Agricultural Economics*, Volume Volume 3, pp. 2857–2908. Elsevier.

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

Dixit, A. and R. Pindyck (1994). *Investment under uncertainty.* Princeton University Press.

Dupas, P. (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica 82*(1), 197–228.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on

moral hazard in health insurance. *American Economic Review 103*(1), 178–219.
Fafchamps, M. (1993). Sequential labor decisions under uncertainty: An estimable household model of west-african farmers. *Econometrica*, 1173–1197.
Fischer, G., D. Karlan, M. McConnell, and P. Raffler (2014). To charge or not to charge: Evidence from a health products experiment in uganda. *Working Paper*.
Foster, A. D. and M. R. Rosenzweig (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 1176–1209.
Giné, X. and D. Yang (2009). Insurance, credit, and technology adoption: Field experimental evidence from malawi. *Journal of Development Economics 89*(1), 1–11.
Goldman, D. P., G. F. Joyce, and Y. Zheng (2007). Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. *Journal of the American Medical Association 298*(1), 61–69.
Heckman, J. J. and S. Navarro (2005). Empirical estimates of option values of education and information sets in a dynamic sequential choice model. *Working paper*.
Heckman, J. J. and S. Navarro (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics 136*(2), 341–396.
Heckman, J. J., D. Schmierer, and S. Urzua (2010). Testing the correlated random coefficient model. *Journal of Econometrics 158*(2), 177–203.
Jack, B. K. (2013). Private information and the allocation of land use subsidies in malawi. *American Economic Journal: Applied Economics 5*, 113–35.
Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics 129*(2), 597–652.
Karlan, D. and J. Zinman (2009). Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica 77*(6), 1993–2008.
Magnan, N., T. J. Lybbert, R. Mrabet, and A. Fadlaoui (2011). The quasi-option value of delayed input use under catastrophic drought risk: the case of no-till in morocco. *American Journal of Agricultural Economics 93*(2), 498–504.
Mahajan, A. and A. Tarozzi (2011). Time inconsistency, expectations and technology adoption: The case of insecticide treated nets. *Working Paper*.
Pindyck, R. (1993). Investments of uncertain cost. *Journal of Financial Economics 34*(1), 53–76.
Saka, A., W. Bunderson, O. Itimu, H. Phombeya, and Y. Mbekeani (1994). The effects of acacia albida on soils and maize grain yields under smallholder farm conditions in Malawi. *Forest Ecology and Management 64*, 217–230.
Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics 4*(1), 49–84.
Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica 79*(1), 159–209.

Appendix to *Technology adoption under uncertainty*

A.1 Conceptual model

This appendix includes the formal proof of Propositions 1 through 4 in the main text. We start by characterizing agents' decisions and types in a more formal way.

A.1.1 Expected value of take-up

The expected net benefit of take-up (which appears in the take-up decision inequality, equation (1) in the main text) can be rewritten as

$$\mathbb{E}_{F_1|F_0}\max(R-F_0-F_1,0) = \Pr(R-F_0-F_1>0|F_0)\times\left[R-F_0-\mathbb{E}_{F_1|F_0}(F_1|R-F_0-F_1>0)\right],$$

where $\Pr(R-F_0-F_1>0|F_0)$ indicates the type-specific (i.e., conditional on F_0) probability of follow-through and $R-F_0-\mathbb{E}_{F_1|F_0}(F_1|R-F_0-F_1>0)$ is the net benefit, conditional on follow-through.

A.1.2 Adoption types

Under the distributional assumptions stated in the main text:

- $F_0 \perp F_1$,
- F_1 takes one of two values: $F_1=\{f_L,f_H\}$, with $f_L<f_H$, and $\mathbb{E}_{F_1}(F_1)=p_Lf_L+(1-p_L)f_H$, where p_L is the probability that F_1 takes the value f_L, and
- F_0 is continuously distributed across agents with cumulative distribution function $\mathrm{G}_0(.)$,

we can classify agents in three follow-through types: non-adopters, contingent adopters and always adopters.

Never follow-through types Never follow-through types are characterized by the condition on F_0,

$$R - F_0 < f_L \tag{1}$$

such that even when the realization of F_1 is low (f_L), their net benefit of follow-through is negative. The share of never follow-through types is given by $1 - \mathrm{G}_1(R - f_L)$. Their probability of follow-through is always 0 and so is their expected private benefit. Never follow-through types take-up only if $c - A > 0$, or if the subsidy exceeds the cost of take-up. Note that even when they take-up (purchase the technology), they never follow-through.

Contingent follow-through types Contingent follow-through types are characterized by the condition

$$f_L < R - F_0 < f_H. \tag{2}$$

Contingent follow-through types follow-through when the realization of F_1 is f_L, but not when the realization is f_H. The share of contingent follow-through types is given by $\mathrm{G}_0(R - f_L) - \mathrm{G}_0(R - f_H)$, with expected private benefit given by

$$\mathbb{E}_{F_1}\left[\max(R - F_0 - F_1, 0) | R - F_0 - F_1 > 0\right] = p_L\left(R - F_0 - f_L\right)$$

where p_L is their probability of follow-through. The take-up decision of these agents is characterized by condition $F_0 \leq R - f_L - \frac{c-A}{\delta p_L}$.

Always follow-through types Always follow-through types are characterized by the condition

$$f_H < R - F_0. \tag{3}$$

Hence, they follow-through whether the draw of F_1 is f_L or f_H: $\Pr(R - F_0 - F_1 > 0 | F_0) = 1$. The share of always follow-through types is given by $\mathrm{G}_0(R - f_H)$, and their private benefit given by

$$\mathbb{E}_{F_1}\left[\max(R - F_0 - F_1, 0) | R - F_0 - F_1 > 0\right] = R - F_0 - \mathbb{E}(F_1).$$

They take-up only if $F_0 < R - \mathbb{E}(F_1) - \frac{c-A}{\delta}$.

A.1.2.1 Selection and follow-through

Conditions (1), (2), and (3) determine thresholds over the support of F_0 that delimit the shares of always adopters, contingent adopters and never adopters for a given distribution of F_0. Figure 1 in the main text illustrates these thresholds on the probability density function of F_0, $\mathrm{g}_0(F_0)$. Note that the bell shaped distribution for F_0 shown in Figure 1 is not a necessary assumption of the model, and is used only to visualize the shares of each agent type as the area under the curve delimited horizontally by the thresholds in gray: $R - f_H$ and $R - f_L$.

The thresholds in black correspond to the take-up decision for each agent type. The take-up threshold for contingent follow-through types, $R - f_L - \frac{c-A}{\delta p_L}$, is always to the right of the threshold that separates contingent follow-through types from never follow-through types provided that the subsidy, A, is less than or equal to the total cost of the technology, c. Hence, the bigger the subsidy, A, the bigger the share who take-up, but follow-through only if $F_1 = f_L$. The take-up threshold for always follow-through types, $R - \mathbb{E}(F_1) - \frac{c-A}{\delta}$, may be to the left or to the right of the threshold, $R - f_H$, which defines the group of always follow-through types. If $\frac{c-A}{\delta} \leq f_H - \mathbb{E}(F_1)$, all always follow-through types will take-up. However, if $\frac{c-A}{\delta} > f_H - \mathbb{E}(F_1)$, a bigger subsidy may increase take-up among always follow-through types.

In sum, the subsidy A affects follow-through rates conditional on take-up by determining the shares of always follow-through types and contingent follow-through types that take up. When the subsidy is small, such that $A < c - \delta(f_H - \mathbb{E}(F_1))$, not all always follow-through types take-up. When the subsidy is between $c - \delta(f_H - \mathbb{E}(F_1))$ and c all always follow-through types take-up, but just a fraction of contingent follow-through types take-up. For subsidies larger than c, all always follow-through types, all contingent follow-through types and some never follow-through types take-up.

Proposition 1 *Follow-through conditional on take-up increases as a function of take-up cost, i.e. there is a screening effect of the take-up cost.*

Conditional follow-through types are the population of interest for understanding the relationship between uncertainty and technology adoption: they constitute the only group whose follow-through decision is affected by the shock realization. The share of conditional follow-through types who take-up is given by

$$\frac{p_L \left[\mathrm{G}_0 \left(R - f_L - \frac{c-A}{\delta p_L} \right) - \mathrm{G}_0(R - f_H) \right] + \mathrm{G}_0 \left(R - f_H \right)}{\mathrm{G}_0 \left(R - f_L - \frac{c-A}{\delta p_L} \right)} \tag{4}$$

if $\frac{c-A}{\delta} < f_H - \mathbb{E}(F_1)$ and is 100 percent if $\frac{c-A}{\delta} \geq f_H - \mathbb{E}(F_1)$. These two expressions show how follow-through depends on A through the take-up decision of the different types of agents: the larger the subsidy, A, the larger the share of contingent follow-through types that take-up, reducing the overall rate of follow-through among those who take-up.

Proposition 2 *An increase in uncertainty reduces follow-through conditional on take-up.*

Note that in expression (4), an increase in the spread of F_1 (distance between f_H and f_L) results in a bigger increase in the denominator than in the numerator, since part of the numerator is multiplied by p_L, which is a number between 0 and 1. Hence, uncertainty worsens follow-through conditional on take-up.

Proposition 3 *An increase in uncertainty weakens the relationship between take-up cost and conditional follow-through shown in Proposition 1.*

Uncertainty increases the share of contingent follow-through types. This is easy to see since the share of contingent follow-through types is determined by the probability mass over the support of F_0 between $R - f_H$ and $R - f_L$. The greater the spread of F_1, the bigger the share of contingent adopters, and the less the take-up decision predicts follow-through. In the extreme case of no uncertainty, there are no contingent follow-through types ($f_H = f_L$) and all we have is either always or never follow-through types. In this case, A increases take-up among always follow-through types, but does not lower follow-through conditional on take-up unless agents are paid to take-up the technology ($A > c$).

A.1.2.2 Option value of the contract

The option value associated with the take-up decision when agents are free to follow-through or not at time 1, i.e. under limited liability, is given by

$$OV(F_0) = \mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) - \max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right) \tag{5}$$

with $\mathbb{E}_{F_1}(R - F_0 - F_1) = R - F_0 - \mathbb{E}(F_1)$, where $\max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right)$ represents the expected profit associated with making the follow-through decision at time 0 and sticking to it, or the value of the static decision. Note that for never follow-through types, the decision to not follow-through does not change with new information. Hence, $\mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) = \max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right) = 0$. Similarly, always always follow-through types' decision does not change with new information. Hence, $\mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) = \max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right) = R - F_0 - \mathbb{E}(F_1)$. Therefore, the only group with a positive option value is contingent follow-through types. For them, $\mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) > \max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right)\mathbb{E}_{F_1} \max(R - F_0 - F_1, 0) > \max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right)$, since

$$\mathbb{E}_{F_1}(R - F_0 - F_1) = R - F_0 - \mathbb{E}(F_1).$$

The share, $G_0(R - \mathbb{E}(F_1)) - G_0(R - F_H)$ would take-up under a contract with commitment (i.e., a static decision where take-up and follow-through decisions are made simultaneously at time 0) since their expected benefit under commitment, $R - F_0 - \mathbb{E}(F_1)$, is greater than zero. The share $G_0(R - F_L) - G_0(R - \mathbb{E}(F_1))$ would only take-up in the contract without commitment, since their expected benefit under commitment, $R - F_0 - \mathbb{E}(F_1$, is less than zero. Hence, for contingent follow-through types,

$$\max\left(\mathbb{E}_{F_1}(R - F_0 - F_1), 0\right) = \begin{cases} R - F_0 - F_1 & \text{if} \quad F_0 < R - \mathbb{E}(F_1) \\ 0 & \text{if} \quad F_0 > R - \mathbb{E}(F_1) \end{cases}$$

From the definition in (5), it follows that the option value for contingent follow-through types with $F_0 < R - \mathbb{E}(F_1)$ is given by

$$\begin{aligned} p_L\left(R - F_0 - f_L\right) - &\quad \left(R - F_0 - \mathbb{E}(F_1)\right) \\ &= \quad [p_L - 1]\left(R - F_0\right) + (1 - p_L) f_H \\ &= \quad (1 - p_L)\left(f_H + F_0 - R\right); \end{aligned}$$

while the option value for contingent follow-through types with $F_0 > R - \mathbb{E}(F_1)$ is equal to their expected private benefit without commitment: $p_L\,(R - F_0 - f_L)$.

In summary, the option value as a function of F_0 is given by

$$OV(F_0) = \begin{cases} 0 & \text{if} \quad F_0 > R - f_L \\ p_L\,(R - F_0 - f_L) & \text{if} \quad R - \mathbb{E}(F_1) < F_0 \leq R - f_L \\ (1 - p_L)\,(f_H + F_0 - R) & \text{if} \quad R - f_H < F_0 \leq R - \mathbb{E}(F_1) \\ 0 & \text{if} \quad F_0 \leq R - f_H \end{cases}$$

Proposition 4

The option value associated with take-up is increasing in uncertainty, which results in higher take-up at all take-up cost levels.

For a given agent with $F_0 = f_0$, option value increases with uncertainty. As uncertainty increases (the distance between f_H and f_L), so does the likelihood that $R - f_H < f_0 \leq R - f_L$, which in turn increases the likelihood that the agent becomes a contingent follow-through type. Hence, as uncertainty increases, the share of agents with a positive option value from take-up also increases. As expected, the option value has an asymmetric relationship with the upper and lower bounds of the shock distribution. One can increase the option value indefinitely by lowering f_L (which is equivalent to increasing the realization of the positive shock, since f_L enters as a cost in the profit function). However, lowering f_H leads to an increase in the option value up to the point where $R - \mathbb{E}(F_1) < f_0$; beyond this, the option value remains constant and equal to $p_L\,(R - F_0 - f_L)$, which is equal to the expected private benefit of the contract to contingent follow through types.

As a function of R, the option value for a given individual with $F_0 = f_0$ is zero up to the point where $R - f_L$ is larger than f_0. Beyond this, the agent becomes a contingent follow-through type and the option value is increasing with R up to $R = f_0 + \mathbb{E}(F_1)$, where it peaks and then falls up to $R = f_0 + f_H$. After this, the option value becomes 0 again since the value of R is large enough to guarantee follow-through.

A.2 Estimation

The estimation of the model outlined in Section 5 in the main text is done via simulated maximum likelihood.[1] This appendix details the estimation procedure used to recover the structural parameters.

A.2.1 Additional parameters

Our field experiment design included two additional treatment arms in addition to the ones described in Section 5.3: a "surprise reward treatment" group and a monitoring group. In the structural estimation, we modify the profit function to account for the variation in choices that these treatment arms introduce.

Surprise reward treatment Half of the farmers who attended training (52.5 percent) were assigned to a "surprise reward treatment" and did not learn about the threshold reward for follow-through (≥ 35 trees) until after their decision to take-up was made. As explained in Section 3.2, this treatment arm allows us to explore whether liquidity constraints explain the absence of selection effects in the data. In order to keep track of the information differences at the time of take-up in the estimation, we allow for these individuals to have a separate component in the profit from planting any positive amount of trees (a constant "surprise treatment" effect, α_S). If these individuals had identical beliefs about the costs and benefits derived from the trees (which in practice means that random parameters F_0, F_1 and T were drawn from the same distribution as those in the standard treatment, who learned about the reward before choosing to take-up), the surprise treatment effect would be zero. However, we observe reduced form evidence that there was an expectation of a higher profit among those who did not know about the reward before taking up: their take-up rate is higher than the rate among farmers who received a reward of zero. The average take-up among those in the surprise reward treatment was approximately equal to the the take-up rate of farmers in the standard treatment who drew a reward of ZMK 40,000 before they made their take-up decision.[2] Hence, in our estimation, the surprise treatment is left unrestricted and is estimated to be 91.79 (s.e. 8.11) in the main model and 54.42 (s.e. 10.235) in the model with a mean shift in F. Note that this latter coefficient is close to the reduced form effect.

Monitoring group A small share of the program participants, 15.8 percent, were randomly selected to receive regular visits to monitor tree-related activities, which allows us to more closely observe time use. This group experienced higher follow-through rates than farmers who were not assigned to the monitoring group,. Though the treatment was not designed to have an impact and the monitors were explicitly told not to communicate information about tree cultivation to the farmers, monitoring may have influenced farmers in

[1] See Train (2009).

[2] This calculation is performed from the results of a linear regression of take-up on the reward among those who had knowledge of the reward before deciding to take up.

a number of ways. For example, monitoring could have increased the subjective value of the trees by making them seem "more important" or decreased the cost of caring for them by periodically reminding individuals of their location and commitment. Farmers were not aware that they would be monitored when they made their take-up decision. In order to account for the observed effect of monitoring in the estimation, we allow the profit of those in the monitoring group to have a separate component that takes the value of zero if no trees are cultivated and of α_M when any positive number of trees is cultivated. This parameter is estimated to be 238.40 (s.e. 36.844) in the main model and 229.53 (s.e. 37.22) in the model with a mean shift in F.

A.2.2 Objective Function details under Simulated Maximum Likelihood

We use simulation methods to evaluate the objective function, equation (8) (from the main text), for any given value of the parameters. We use simulated maximum likelihood because there are several quantities in our objective function that do not have a closed form expression.

As is usually done in random parameter models, we integrate away the unobserved random parameters when writing the analytic probabilities for each outcome. These integrals, once more, do not have a closed form solution. Hence, we use numerical integration to write the probability of choosing N trees conditional on parameters $\mu_F, \sigma_{F_0}, \sigma_{F_1}, \mu_T, \sigma_T, \alpha_S,$ and α_M. Before writing the expression for the simulated probabilities of choosing N trees, we note one more aspect of our estimation strategy.

When using simulation methods to estimate discrete choice models with random parameters, numerical integrals are used to approximate theoretical probabilities. This often results in a stepwise as opposed to smooth objective function, since small probabilities are hard to approximate numerically and can be very noisy. In order to smooth the kinks in our objective function, we add an extreme value distributed error term at the end of the profit function. This allows us to compute probabilities between 0 and 1 for each draw of the random parameters, which results in a smoother objective function. Monte Carlo simulations suggest this method will not introduce bias our results provided that we choose a relatively small scale parameter, λ, which we refer to as smoothing factor. In the estimation we use a smoothing factor of 0.5.

Thus, using Train (2009) notation, the simulated probabilities of choosing N trees at $t = 1$ are

$$\check{P}_i(N^* = n|\theta) = \frac{1}{K}\sum_{k=1}^{K} \frac{\exp\left(\frac{1}{\lambda}\Pi(n|F_{0k}, F_{1k}, T_k, R_i)\right)}{\sum_{j=0}^{50}\exp\left(\frac{1}{\lambda}\Pi(j|F_{0k}, F_{1k}, T_k, R_i)\right)}$$

where k indexes each draw of the full random parameter vector, (F_{0k}, F_{1k}, T_k), given the vector of parameters $\theta = (\mu_F, \sigma_{F_0}, \sigma_{F_1}, \mu_T, \sigma_T, \alpha_S, \alpha_M)$, and farmer-specific treatments A_i and R_i.

Similarly, the simulated probability of take-up at $t = 0$ is given by

$$\check{P}_i(TakeUp|\theta) = \frac{1}{K}\sum_{k=1}^{K} \mathbf{1}\left(A_i - c + \delta\check{\mathbb{E}}\left[\max_N \Pi(N|T_k, F_{0k}, F_1, T_i, R_i)|\, F_{0k}, T_k\right] > 0\right) \tag{6}$$

where k indexes each draw of the partial random parameter vector, (F_{0k}, T_k), given the vector of parameters $\theta = (\mu_F, \sigma_{F_0}, \sigma_{F_1}, \mu_T, \sigma_T, \alpha_S, \alpha_M)$, and farmer-specific treatments A_i and R_i. Note that the expected profit conditional of random variables F_0 and T, and observed treatments A and R also involves an integral without a closed form solution. We therefore use the simulated version of it in expression (6). More specifically,

$$\check{\mathbb{E}}\left[\max_N \Pi(N|T_k, F_{0k}, F_1, T_i, R_i)|\, F_{0k}, T_k\right] = \frac{1}{M}\sum_{m=1}^{M} \max_N \quad \Pi(N|F_{0k}, F_{1m}, T_k, R_i, A_i) \tag{7}$$

where m indexes each of M draws from a normal distribution with mean F_{0k} and variance given by $\sigma^2_{F_1}$.

For estimation purposes, we use $K = 1500$ and $M = 100$. Each of the k draws are independent across observations. However, the M draws used in (7) are kept constant across observations. This reduces our computing power substantially without affecting the independence assumptions across observations (note that (7) is conditioned on F_{0k} and T_k, which are drawn independently for each farmer).

A.2.3 Maximization algorithm

In order to guarantee that the point estimates correspond to the global maximum of the likelihood function, we first conducted a grid search that would inform our starting values for the numerical maximization. The grid search was conducted over 80 thousand different combinations of the parameters and, to minimize computing time, was conducted with a lower value of K and M (400 and 50 respectively).

In addition, we conducted a three-stage recursive maximization (minimization of the negative likelihood) where in each stage we maximized the simulated likelihood along a subset of the parameter vector holding the rest constant. This method worked better than the single step maximization in Monte Carlo simulations. The subsets of parameters in each of the three stages were (μ_T, σ_T, ρ), $(\sigma_{F0}, \sigma_{F1})$, and $(\mu_F, \alpha_M, \alpha_S, \mu_{shift})$ respectively.[3] The three stages were repeated sequentially until a convergence criterion involving changes in the parameter values was reached.[4] In a final stage, we used the resulting parameter estimates as starting values in a single step numerical maximization. This last step yielded small changes in the parameter values (the largest change was less than 6 percent and corresponded to the

[3] μ_{shift} corresponds to the common uniform shock in the mean shifter model discussed in Sections 5 and 6.

[4] The convergence criterion we used was that the square sum of differences between the new parameters and the starting values (the estimated parameters from the last optimization round) was less than 0.0001. The number of iterations was very robust to the critical value chosen and never reached more than four iterations.

monitoring parameter, α_M; the second largest change was of 4 percent and corresponded to the standard deviation of F_0, σ_{F0}). Appendix Table A.2.1 shows the sensitivity of our three-stage results to starting values slightly above, slightly below and at the parameter values that maximized the likelihood in our grid search.

A.2.4 Standard errors

Standard errors were computed using the variance of the numerically approximated scores, which should converge to the negative of the Hessian in the limit provided that the point estimates are the argmax of the log-likelihood function (Train 2009). We chose this method instead of the numerical Hessian because it allowed us to choose the size of the step (h) when calculating the numerical score. Simulated methods often result in "roughness" of the likelihood function, which, in our case, led to a non-positive definite numerical Hessian.[5] In order to verify that we were at a (local) minimum, we plotted the likelihood to verify its curvature along each parameter, one at a time. Appendix Table A.2.2 shows the sensitivity of our standard errors to different values of h. We chose the value of h that led to the smallest gradient.

[5]The default numerical gradient calculation also led to gradient components that far from zero. In contrast, all elements of the numerical gradient we "manually" calculated were very close to zero.

Table A.2.1: Parameter sensitivity to starting values

Panel A. Tuning parameter sets

Estimation Tuning Parameters

Name	Lambda	DiffMinChange	k	m
Set A	0.5	0.1/0.5/2	1500	100
Set B	0.5	0.05/0.05/0.5	1500	100
Set C	2	0.1/0.5/2	1500	100
Set D	0.5	0.1/0.5/2	2500	200

Panel B. Choice set of initial parameters

List of Initial Values of Parameters

μ_T	σ_T	σ_{F0}	σ_{F1}	ϱ	μ_F	α_s	α_m	μ_{Fs}
2	0.5	450	50	0.5	0	0	-100	0
3.2	1.2	500	100	0.7	100			20
3.8	1.8			0.8				50

Panel C. List of runs by initial values and parameters

	Description	Tuning Set	μ_T	σ_T	σ_{F0}	σ_{F1}	ϱ	μ_F	α_s	α_m	μ_{Fs}
			Initial Values of Parameters								
1	Values close to gridsearch	Set A	3.2	1.2	450	100	0.7	100	0	-100	20
2	Values below gridsearch optimum	Set A	2	0.5	450	50	0.5	0	0	-100	0
3	Values above gridsearch optimum	Set A	3.8	1.8	500	100	0.8	100	0	-100	50
4	Below for muT, above for rho	Set A	2	0.5	450	50	0.8	100	0	-100	20
5	Above for muT, below for rho	Set A	3.8	1.8	450	100	0.5	100	0	-100	20
6	All values close to gridsearch except for rho	Set A	3.2	1.2	450	100	0.8	100	0	-100	20
7	All values close to gridsearch except for muT, sdT	Set A	3.8	1.8	450	100	0.7	100	0	-100	20
8	All values close to gridsearch except for muF	Set A	3.2	1.2	450	100	0.7	0	0	-100	20
9	Changing DiffMinChange	Set B	3.2	1.2	450	100	0.7	100	0	-100	20
10	Changing Lambda	Set C	3.2	1.2	450	100	0.7	100	0	-100	20
11	Changing k	Set D	3.2	1.2	450	100	0.7	100	0	-100	20

Panel D. Results without Fshifter

	Description	Log-likelihood	μ_T	σ_T	σ_{F0}	σ_{F1}	ϱ	μ_F	α_s	α_m	μ_{Fs}
			Values of Structural Parameters at Termination								
1C	Values close to gridsearch	11151.17	3.292	1.262	291.12	190.22	0.658	73.54	-79.56	-134.08	-
2C	Values below gridsearch optimum	11160.82	3.132	1.213	188.23	135.91	0.614	60.20	-73.04	-107.22	-
3C	Values above gridsearch optimum	11144.84	3.535	1.399	312.11	202.28	0.803	108.10	-94.64	-224.94	-
4C	Below for muT, above for rho	11168.35	3.235	1.203	136.22	138.87	0.723	84.73	-35.43	-109.05	-
5C	Above for muT, below for rho	11152.62	3.395	1.330	297.10	202.58	0.600	99.11	-56.53	-150.44	-
6C	All values close to gridsearch except for rho	11145.23	3.534	1.364	299.24	200.99	0.775	103.28	-96.51	-221.31	-
7C	All values close to gridsearch except for muT, sdT	11147.79	3.426	1.340	297.28	199.67	0.729	97.16	-93.31	-184.77	-
8C	All values close to gridsearch except for muF	11153.41	3.269	1.225	276.41	191.96	0.684	73.60	-75.01	-147.78	-
9C	Changing DiffMinChange	11145.84	3.416	1.355	288.73	226.57	0.653	107.94	-24.03	-253.77	-
10C	Changing Lambda	# 8295.695	3.117	1.267	300.38	201.87	0.668	75.26	-81.22	-262.02	-
11C	Changing k	# 11147.7136	3.434	1.388	296.62	203.85	0.631	84.61	-57.01	-129.97	-

Panel E. Results with Fshifter

	Description	Log-likelihood	Values of Structural Parameters at Termination								
			μ_T	σ_T	σ_{F0}	σ_{F1}	ϱ	μ_F	α_s	α_m	μ_{Fs}
1D	Values close to gridsearch	11153.662	3.469	1.389	279.40	181.33	0.658	51.51	-38.56	-113.99	31.75
2D	Values below gridsearch optimum	11159.819	2.977	1.169	214.61	155.62	0.365	44.00	-36.97	-107.84	37.94
3D	Values above gridsearch optimum	11141.838	3.550	1.389	301.93	195.60	0.823	76.08	-55.03	-217.74	53.61
4D	Below for muT, above for rho	11150.377	3.293	1.268	174.11	134.94	0.753	57.42	-47.36	-158.37	30.79
5D	Above for muT, below for rho	11146.150	3.319	1.256	287.15	190.26	0.592	58.72	-48.54	-222.04	48.25
6D	All values close to gridsearch except for rho	11143.875	3.510	1.365	289.78	198.01	0.772	63.46	-40.43	-189.88	47.32
7D	All values close to gridsearch except for muT, sdT	11145.224	3.400	1.320	288.39	194.84	0.721	71.01	-37.69	-231.67	52.93
8D	All values close to gridsearch except for muF	11145.255	3.381	1.332	225.21	158.53	0.661	63.87	-65.28	-167.64	31.47
9D	Changing DiffMinChange	11150.482	3.255	1.155	267.12	202.41	0.622	85.17	-37.48	-220.85	21.86
10D	Changing Lambda	# 8296.7806	3.052	1.203	293.91	197.92	0.669	55.51	-39.43	-252.52	13.73
11D	Changing k	# 11145.1056	3.444	1.390	289.29	199.59	0.629	68.82	-35.18	-173.34	45.78

Panel F. Final minimization using above results

Description	Log-likelihood	Values of Structural Parameters at Termination								
		μ_T	σ_T	σ_{F0}	σ_{F1}	ϱ	μ_F	α_s	α_m	μ_{Fs}
Initial values are 3C results, gradient free algorithm	11142.064	3.539	1.401	307.87	211.42	0.818	107.58	-91.79	-238.40	-
Initial values are 3D results, gradient free algorithm	11138.996	3.579	1.392	290.06	193.05	0.835	74.48	-54.42	-229.53	53.29

Notes: This table list the set of minimization attempts conducted based on grid search results and model exploration. Panels A and B list the choice set of tuning parameters and initial values that are used. Panel C describes the tuning parameters and initial values used for both models. Panels D and E follow the same order as Panel C, listing the results and negative log-likelihood values for each run with and without the F shifter. Panel F takes the best run from Panels D and E and reports the results of an additional minimum search using a gradient free method. # indicates that the log-likelihood cannot be compared to other values.

Table A.2.2: Alternative standard error calculations

Additive h	Values of k										
h = k	0.09	0.07	0.05	0.03	0.01	0.009	0.007	0.005	0.003	0.001	0.0005
SSE(gradient)	0.00020	0.00007	0.00003	0.00030	0.00116	0.00094	0.00047	0.00015	0.00012	0.00123	0.00087
Standard Errors											
μ_T	0.0788	0.0569	0.0536	0.0446	0.0361	0.0363	0.0409	0.0555	0.0742	0.0783	0.0745
σ_T	0.0805	0.0661	0.0720	0.0441	0.0353	0.0358	0.0399	0.0499	0.0585	0.0568	0.0537
ϱ	0.0725	0.0656	0.0578	0.0549	0.0424	0.0407	0.0380	0.0346	0.0293	0.0205	0.0132
σ_{F0}	95.341	93.278	82.250	77.089	58.658	57.753	60.026	57.385	47.803	37.121	27.795
σ_{F1}	50.621	49.953	44.387	42.643	33.717	32.925	33.923	32.853	26.950	21.558	16.626
μ_F	13.083	11.822	10.323	7.398	4.578	4.268	3.875	3.441	3.319	1.129	0.587
α_s	18.244	16.222	13.914	9.604	6.066	5.725	5.054	4.715	4.839	2.945	1.478
α_m	74.432	73.887	69.830	70.173	67.778	67.349	69.312	69.247	64.716	61.108	59.251

Multiplicative h	Values of k										
h = x * k	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001	0.0005	0.0002	0.0001
SSE(gradient)	0.17430	0.00042	0.00099	0.00035	0.00146	0.00449	0.00667	0.00803	0.00840	0.00953	0.01186
Standard Errors											
μ_T	0.1847	0.1451	0.1006	0.0584	0.0438	0.0366	0.0252	0.0219	0.0209	0.0205	0.0203
σ_T	0.0741	0.0420	0.0287	0.0216	0.0170	0.0161	0.0144	0.0137	0.0136	0.0134	0.0129
ϱ	0.1285	0.1066	0.0791	0.0492	0.0374	0.0348	0.0257	0.0213	0.0172	0.0140	0.0111
σ_{F0}	12.250	111.716	114.417	98.676	80.514	56.321	37.237	28.183	22.570	17.511	14.339
σ_{F1}	6.171	61.562	63.773	54.957	45.919	33.423	20.956	16.169	13.436	10.346	7.947
μ_F	27.740	27.335	23.773	19.440	17.004	15.048	11.832	9.402	6.050	3.819	2.419
α_s	40.490	46.246	42.654	38.907	33.646	27.268	20.378	15.132	9.517	8.783	6.360
α_m	78.068	111.782	110.726	95.144	76.055	65.208	57.078	55.481	54.674	54.310	54.300

Additive h	Values of k										
h = k	0.09	0.07	0.05	0.03	0.01	0.009	0.007	0.005	0.003	0.001	0.0005
SSE(gradient)	0.00005	0.00005	0.00014	0.00037	0.00400	0.00425	0.00371	0.00215	0.00054	0.00005	0.00018
Standard Errors											
μ_T	0.0706	0.0682	0.0570	0.0443	0.0371	0.0372	0.0412	0.0486	0.0597	0.0778	0.0730
σ_T	0.0747	0.0799	0.0604	0.0553	0.0352	0.0345	0.0358	0.0379	0.0422	0.0533	0.0499
ϱ	0.0730	0.0645	0.0561	0.0475	0.0303	0.0299	0.0304	0.0301	0.0283	0.0191	0.0156
σ_{F0}	84.622	83.916	81.794	77.381	59.699	56.483	51.711	56.478	51.760	38.025	25.406
σ_{F1}	45.427	45.133	43.960	41.906	34.570	33.396	30.393	32.730	28.430	20.956	13.530
μ_F	15.470	12.581	9.894	6.373	4.062	3.654	2.976	2.176	1.549	0.534	0.296
α_s	20.470	17.634	13.734	9.510	5.561	5.035	4.021	3.800	2.566	1.402	2.400
α_m	74.444	74.339	73.968	75.463	69.887	69.394	68.504	71.408	70.636	65.708	61.092
μ_{Fs}	26.761	25.478	24.328	23.247	22.047	21.828	21.563	21.827	21.527	20.989	20.369

Panel D. F shifter standard errors -- multiplicative h

Multiplicative h	Values of k										
h = x * k	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001	0.0005	0.0002	0.0001
SSE(gradient)	0.23205	0.00143	0.00059	0.00336	0.00084	0.00107	0.00017	0.00009	0.00006	0.00003	0.00017
Standard Errors											
μ_T	0.1857	0.1565	0.1018	0.0732	0.0496	0.0378	0.0550	0.0758	0.0719	0.0618	0.0519
σ_T	0.0839	0.0464	0.0265	0.0169	0.0148	0.0178	0.0348	0.0505	0.0486	0.0416	0.0344
ϱ	0.1275	0.1045	0.0786	0.0572	0.0429	0.0321	0.0307	0.0267	0.0225	0.0173	0.0132
σ_{F0}	10.866	94.981	92.906	87.104	84.390	76.228	62.849	54.639	55.037	41.367	28.522
σ_{F1}	5.403	61.294	60.337	52.771	47.672	43.032	36.870	32.005	30.577	21.769	14.634
μ_F	41.737	63.408	55.454	43.936	33.010	26.111	15.069	12.822	7.683	7.682	5.593
α_s	44.332	42.804	41.230	35.285	32.598	26.815	18.704	13.748	8.751	6.537	4.262
α_m	69.865	97.904	95.421	92.248	83.544	78.012	71.757	68.865	73.371	66.278	61.556
μ_{Fs}	40.510	59.306	53.847	47.573	38.949	32.736	25.922	25.112	22.604	21.766	20.548

Notes: These tables display potential standard error calculations based on different estimates of the numerical derivative. Dotted boxes indicate standard errors reported in the structural parameters table. Panels A and B report the model with no F shifter, while Panels C and D correspond to the model with the F shifter. The central numerical deriviative is calculated as f(x+h)-f(x-h)/2h. In Panels A and C, the change in x given by h is additive across all variables, so h = k. In Panels B and D, the change in x is a multiplicative with respect to x, so h = x * k. These changes in x refer to the transformation of the parameters to an unbounded space -- all standard errors estimates are calculated using these inputs and the delta method. The calculations are performed on the scores of individual log-likelihoods then summed to estimate the gradient. SSE(gradient) gives the sum of squared errors of this gradient estimate, where the true value is assumed to be zero for the minimizing

A.3 Deterministic tree survival assumption

One of the assumptions in the specification of the farmer's optimization problem is that survival of trees is deterministic, conditional on effort. We allow for the cost of tree cultivation to be quadratic in the number of trees, which would capture increasing marginal costs of tree cultivation arising from increasing marginal opportunity cost of time. Our assumption on deterministic survival can be thought of as a two stage optimization process, where the farmer decides on the optimal number of trees to keep alive first, and then allocates the amount of costly effort that guarantees survival to each of those trees. This assumption is less restrictive than one would think.

The two-stage optimization process is roughly consistent with standard optimization under probabilistic survival with a few restrictions: that the probability of tree survival for a single tree as a function of effort, $p(e)$, (a) is independent across trees; (b) attains 1 at some level of effort, $\tilde{e}$, and (c) is a convex function of effort up to $\tilde{e}$; that is $\lim_{e \to \tilde{e}} p'(e) > 0$. In addition we maintain the standard interior solution assumptions of the profit function: (d) increasing and convex cost of effort, $c(e)$ (i.e. $c'(e) > 0$, $c''(e) > 0$), and (e) diminishing marginal returns to the additional tree. We can denote this last assumption as $g_i > g_{i+1}$, where g_i denotes the marginal benefit of the ith tree that survives. Assumption (c) guarantees that the optimal allocation of effort across two or more trees, given an optimal level of total effort $\bar{e}$, is such that the farmer will allocate $\tilde{e}$ to as many trees as possible up to $k\tilde{e} \leq \bar{e}$. If $k\tilde{e} < \bar{e}$, then only the last tree $(k+1)$ will be allocated the remaining effort, $\bar{e} - k\tilde{e}$, making its survival probability less than one. This optimal allocation of effort is thus consistent with deterministic survival of all trees the farmer cultivates, except for possibly the very last tree.[6]

It could be, however, that no amount of effort guarantees the survival of a given tree: i.e., the probability function reaches a maximum of $p\left(\approx\atop e\right) < 1$ at $\overset{\approx}{e}$. In order to explore whether such a model fits our data better, we simulate farmer's behavior assuming this is the case. We keep the parameters that govern farmers' heterogeneity and shocks from our estimated mean-shift model, and we add probabilistic survival to the argument of the indicator function for reaching the 35-tree threshold.[7] Appendix table A.3.1 replicates the reduced form comparison exercise in Table 3 of the main text under this alternative assumption. For ease of comparison, Panel A shows the reduced form results using the observed data (i.e. is identical to Panel A of Table 4 in the main text). Panel B shows the reduced form results with simulated data under our baseline deterministic tree survival assumption and the estimated parameters of our mean shifter model (i.e. is identical to Panel C of Table 3). Panels C - E implement the same regressions, with simulated data from a model that keeps our estimated parameters constant (Panel B of Table 3), but models tree survival outcomes as stochastic and governed by either a binomial distribution (Panels C and D) or a beta binomial distribution (Panel E).[8] Panel D assumes that the maximum probability of survival, $p(\overset{\approx}{e})$, is 0.98,

[6]The proof behind this optimal distribution of effort across trees consists of showing that there are no interior solutions to the optimization problem where more than one tree is allocated an amount of effort between 0 and $\tilde{e}$. We can prove this by contradiction for the case of two trees. The proof can be easily extended to an unlimited number of trees.

The farmer's maximization problem in the case of two trees is given by

$$\max_{e_1,e_2} \pi(e_1, e_2) = g_1 p(e_1) + g_2 p(e_2) - c(e_1 + e_2)$$

where $g_1 > g_2$ (because of assumption (e)), $p(.)$ meets assumptions (a), (b) and (c), and $c(.)$ meets assumption (d).

For a solution to this problem where both trees receive an amount of effort between 0 and $\tilde{e}$ to exist (i.e. $0 < e_1^* < \tilde{e}$ and $0 < e_2^* < \tilde{e}$), the following condition needs to be satisfied

$$g_1 p'(e_1^*) - c'(e_1^* + e_2^*) = g_2 p'(e_2^*) - c'(e_1^* + e_2^*)$$

which can be simplified to

$$g_1 p'(e_1^*) = g_2 p'(e_2^*) l \tag{8}$$

Because $g_1 > g_2$, and $p''(e) > 0$ for $0 < e < \tilde{e}$, condition (8) requires that $e_1^* < e_2^*$. However, it is easy to see that given a constant total amount of effort, e^*, no optimal distribution of this effort, (e_1^*, e_2^*) will be such that $e_1^* + e_2^* = e^*$ and $e_1^* < e_2^*$ as $g_1 p'(e_1) > g_2 p'(e_2)$ for all $e_1 \leq e_2$. I.e., given a constant amount of total effort, the farmer can always do better reallocating some effort to the tree that has the higher return. Thus, no interior solution exists where more than one tree is receiving an amount of effort less than the minimum amount that guarantees survival, $\tilde{e}$.

[7]Recall that the continuous component of the profit function confounds marginal costs and benefits. Thus we cannot introduce probabilistic survival to the benefit portion, without affecting the cost per-tree, which should remain deterministic.

[8]We keep the estimated parameters under the deterministic survival assumption instead of reestimating them under the stochastic survival assumption due mainly to computing time constraints. Thus, the fit of the model may further improve if we let other parameters adjust instead of keeping them constant. However, the little sensitivity of the reduced form responses we see in Appendix table A.3.1 leads us to believe that

while Panel D assumes that this maximum survival probability is 0.95. We chose relatively high probabilities for the simulation as lower probabilities result eliminate bunching at 35, and thus are inconsistent with what we observe in our data (see Figure A.3.1). The beta binomial distribution in Panel E allows for the maximum probability to vary across farmers according to a beta distribution with parameters 0.57 and 0.37. The purpose of this exercise is to examine whether by relaxing the deterministic survival assumption we can do a better job matching the reduced form results in Panel A than do our main estimates, Panel B.

Overall, we see little improvement when stochasticity is introduced into the tree survival outcomes. The main model performs least well on the relationship between the take-up subsidy and the positive number of trees and zero trees (Panel B, columns 3 and 4). Both models overestimate the effect of the reward on the likelihood of reaching the 35-tree threshold and the number of trees for farmers with any surviving trees (Panel B, columns 6 and 7). The model variants in Panel C and D show no improvement on any of these dimensions, and in some cases worsen the fit. Only Panel E improves on the fit compared to our main model (Panel B), and not by much: the coefficients on the reward for the 35-tree threshold attainment and for positive tree survival are closer to the observed data but qualitative differences remain. Importantly, these improvement come at the expense of a poorer match in other responses that are well-fit by our main model, such as the relationship between the reward and take-up and the relationship between the reward and zero-trees cultivated (columns 5 and 8).

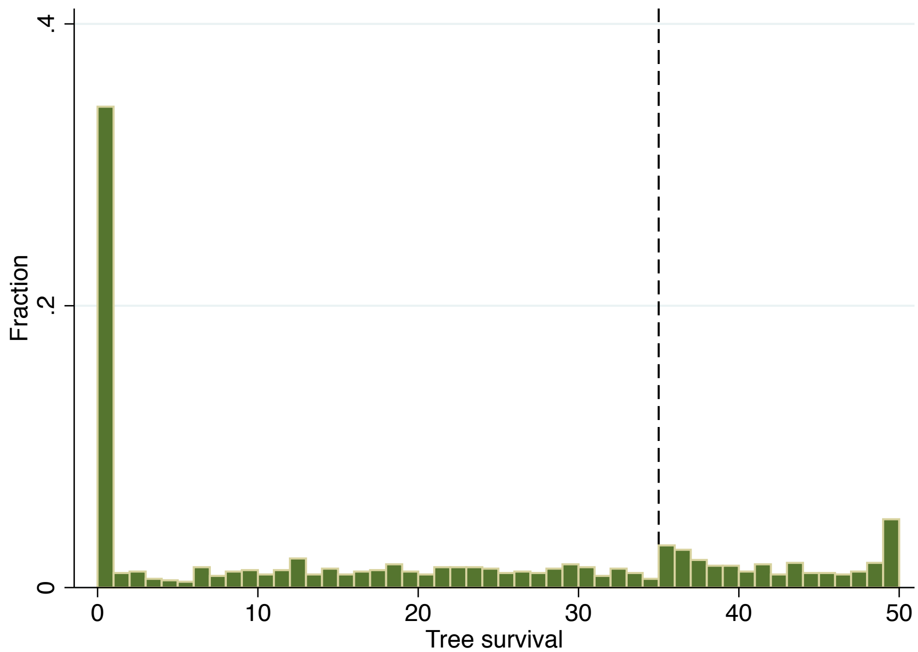
we would not gain much in terms of fit by reestimating the model under the stochastic survival assumption.

Table A.3.1: Stochastic tree survival

	(1) Take-up	(2) 35-tree threshold	(3) # trees | # trees>0	(4) 1.(zero trees)		(5) Take-up	(6) 35-tree threshold	(7) # trees | # trees>0	(8) 1.(zero trees)
Panel A. Observed Data (Repeats Panel A in Reduced Form Table)									
Take-up subsidy	0.022***	-0.004	-0.229	-0.003	Reward	0.001*	0.001***	0.044***	-0.001***
	(0.005)	(0.004)	(0.200)	(0.005)		(0.000)	(0.000)	(0.013)	(0.000)
Observations	1,314	1,092	701	1,092		624	1,092	701	1,092
R-squared	0.071	0.002	0.005	0.001		0.006	0.018	0.022	0.019
Panel B. Mean Shift and No Stochastic Survival (Repeats Panel C in Reduced Form Table)									
Take-up subsidy	0.020***	-0.003	-0.002	0.008**	Reward	0.001*	0.003***	0.094***	-0.001***
	(0.002)	(0.003)	(0.124)	(0.003)		(0.000)	(0.000)	(0.012)	(0.000)
Observations	1,314	1,120	605	1,120		624	1,120	605	1,120
R-squared	0.062	0.001	0.000	0.006		0.006	0.107	0.089	0.013
Panel C. Survival probability = 0.98									
Take-up subsidy	0.020***	-0.002	0.062	0.009**	Reward	0.001*	0.003***	0.105***	-0.001***
	(0.002)	(0.003)	(0.133)	(0.003)		(0.000)	(0.000)	(0.012)	(0.000)
Observations	1,314	1,120	603	1,120		624	1,120	603	1,120
R-squared	0.062	0.000	0.000	0.006		0.006	0.108	0.109	0.012
Panel D. Survival probability = 0.95									
Take-up subsidy	0.020***	-0.002	0.083	0.008**	Reward	0.001*	0.003***	0.101***	-0.001***
	(0.002)	(0.003)	(0.137)	(0.003)		(0.000)	(0.000)	(0.013)	(0.000)
Observations	1,314	1,120	596	1,120		624	1,120	596	1,120
R-squared	0.062	0.000	0.001	0.006		0.006	0.095	0.098	0.012
Panel E. Survival probability distributed beta binomal with mean 0.57 and sd 0.37									
Take-up subsidy	0.022***	-0.001	-0.038	0.006*	Reward	-0.000	0.001***	0.057***	-0.000
	(0.002)	(0.002)	(0.151)	(0.003)		(0.000)	(0.000)	(0.013)	(0.000)
Observations	1,314	1,099	518	1,099		624	1,099	518	1,099
R-squared	0.071	0.000	0.000	0.003		0.000	0.019	0.034	0.002

Notes: This table shows coefficients from regressions of each of four indicator variables (take-up, binary 35-tree threshold, tree survival larger than zero, and no tree survival) on each of our randomized treatments (take-up subsidy and threshold reward) for both non-stochastic and stochastic models. Panel A shows these regression outcomes for the true data. Panel B shows the fit of the structural model by simulating all four outcomes using the model estimates and examining the how much the linear relationships between outcomes and treatments resemble those in Panel A. These panels recreate Panels A and C of the reduced form table in the main body of the paper. Panel C here estimates binomial survival assuming that the probabilty any one tree survives is 0.98. Panel D estimates binomial survival assuming that the probabilty any one tree survives is 0.95. Panel E assumes that the probability any one tree survives is distributed beta binomal with mean 0.57 and standard deviation 0.37, corresponding to an alpha parameter of 0.428 and a beta parameter of 0.318.

Figure A.3.1: Observed tree survival outcomes



Notes: Histogram of tree survival outcomes for all farmers assigned a positive reward for reaching the 35-tree sur
The threshold is shown by the dashed vertical line.

A.4 Confounds and alternative explanations

A.4.1 Liquidity constraints

Our design allows us to address concerns that liquidity constraints complicate the relationship between the take-up decision and farmers' expected payoffs at the time of take-up. We do this in two ways. First, as noted above, all farmers received a training show up fee sufficient to cover the cost of take-up in even the lowest subsidy treatment. Thus, cash-on-hand is unlikely to interfere with take-up. Second, the variation in the timing of the reward for follow-through provides a separate test for self-selection, which does not depend on immediate concerns about liquidity. Specifically, because some farmers were aware of the rewards at the time of the take-up decision and some were not, self selection based on expected payoffs should also incorporate the value of the reward. Importantly, the reward is paid after a year, and thus the response to it at the time of take-up should not be contaminated by immediate liquidity constraints. We see no difference in the response of follow-through to the value of the reward based on the timing of the reward announcement (see Appendix table A.4.1). This provides further support for the conclusion that the new information after take-up plays a substantial role in the follow-through decision.

A.4.2 Psychological channels

Sunk cost, information signaling and crowding out effects

Previous studies of the effect of subsidies on follow-through, via screening on private benefits, have worried about psychological effects associated with the initial price paid for the technology. First, sunk cost effects would cause higher follow-through among adopters who pay for more take-up, because adopters would consider their expenditure at take-up when making their follow-through decision (Ashraf et al. 2010; Berry et al. 2012; Cohen and Dupas 2010). Second, farmers could extract information about the quality of the technology from the NGO's decision to subsidize (Milgrom and Roberts 1986). If higher subsidies accompany better technologies, then farmers in a higher subsidy condition might have a higher follow-through (tree survival). This works in the opposite direction as the sunk cost effect. Third, if paying farmers to take up the technology crowds out their intrinsic motivation for growing trees, we might see higher take-up subsidies leading to lower follow-through (Benabou and Tirole 2003; Bénabou and Tirole 2006). This crowding out effect would work in the same direction as the sunk cost effect. Because we observe no effect of the exogenous variation in take-up subsidies on follow-through, we are able to rule out all three of these explanations.[9] Note that were we to observe a screening effect of the take-up price, we would not be able to distinguish it from these other channels using our design.

[9] That we do not see differences in follow-through across the reward timing conditions provides additional evidence against crowding out of intrinsic motivation associated with selecting in to the program based on expectation of a reward.

House money or experimenter demand effects

The use of "house money" (the show up fee paid to all farmers who joined the training) to purchase the seedlings may have affected farmers' decisions to take-up the technology, confounding our interpretation of selection effects. While we cannot rule out that some farmers' decisions to take-up were influenced by factors other than expected net profits, we can test whether, on average, farmers were more likely to take up because they felt an obligation to do so, or felt reciprocity toward the implementing organization. The high observed rates of take-up in at all levels of the subsidy makes this a potential concern.[10]

We use estimates of the mean shift between take-up and follow-through, μ_{F_S}, which can be interpreted as a change in the payoffs facing all farmers on average between the time of take-up decision and follow-through. The estimated value is positive but small and not significantly different from zero at standard confidence levels. While the sign on the term is consistent with an experimenter demand effect, its magnitude is small compared to the standard deviation of the shocks, σ_{F_1}. Thus, these types of psychological factors that may have driven take-up appear to play a relatively small role, on average, in our setting.

A.4.3 Spillovers

Our experimental treatment arms vary the take-up subsidy at the group level and the threshold reward for tree survival at the individual level. By design, the group level variation is unlikely to lead to spillovers (no two farmer groups are in the same village). The individual variation in the reward level raises the possibility of spillovers that affect both take-up and follow-through.

Effect of reward spillovers on take-up

Even though take-up was decided in a one-on-one interaction with the enumerator and an effort was made to suppress communication until all decisions were made, farmers may still have been able to communicate information about the rewards to those who had not yet taken up. We test for spillovers across reward outcomes by regressing the probability of take-up on the average random reward draw that preceded a farmer's own draw. Regardless of the number of preceding draws we include in the regression, the outcomes of preceding draws have no effect on the probability of take-up.

Effect of reward spillovers on follow-through

Randomization of rewards at the individual level also gives rise to concerns 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

[10] The alternative, not providing a show up fee, would have introduced a more serious confound, in our opinion. As discussed in the previous sub-section on liquidity constraints, the show up fee was necessary to ensure that take-up was not driven by cash-on-hand rather than expected net profits.

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 take-up subsidy shows no significant interaction or clear pattern of coefficients. Second, to investigate the potential for transplanting, we take advantage of a brief survey conducted on all farmers who took up shortly after the end of the rainy season, when most planting would have occurred. 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. Appendix table A.4.2 shows spillover effects of the rewards.

Table A.4.1: Effect of reward timing on tree survival outcomes

	(1) Surprise = 0 Mean/[SD]	(2) Surprise = 1 Mean/[SD]	(3) Reward x Surprise Coef/(SE)
R = 0	11.02	11.32	
	[14.33]	[16.00]	
R = (0,70000)]	14.71	15.87	0.85
	[16.86]	[16.87]	(2.72)
R = (70000,150000]	20.32	21.05	0.43
	[17.99]	[17.48]	(2.66)

Notes: Outcome is tree survival (continuous), conditional on take up. Columns 1 and 2 show means and standard deviations in each reward category, by the reward timing condition. Surprise = 1 indicates that farmers learned about the reward only after the take-up decision. Column 3 reports estimated coefficients and standard errors clustered at the farmer group level for a linear regression of tree survival on reward category interatcted with the surprise reward treatment. We report the coefficient on the interaction term only.

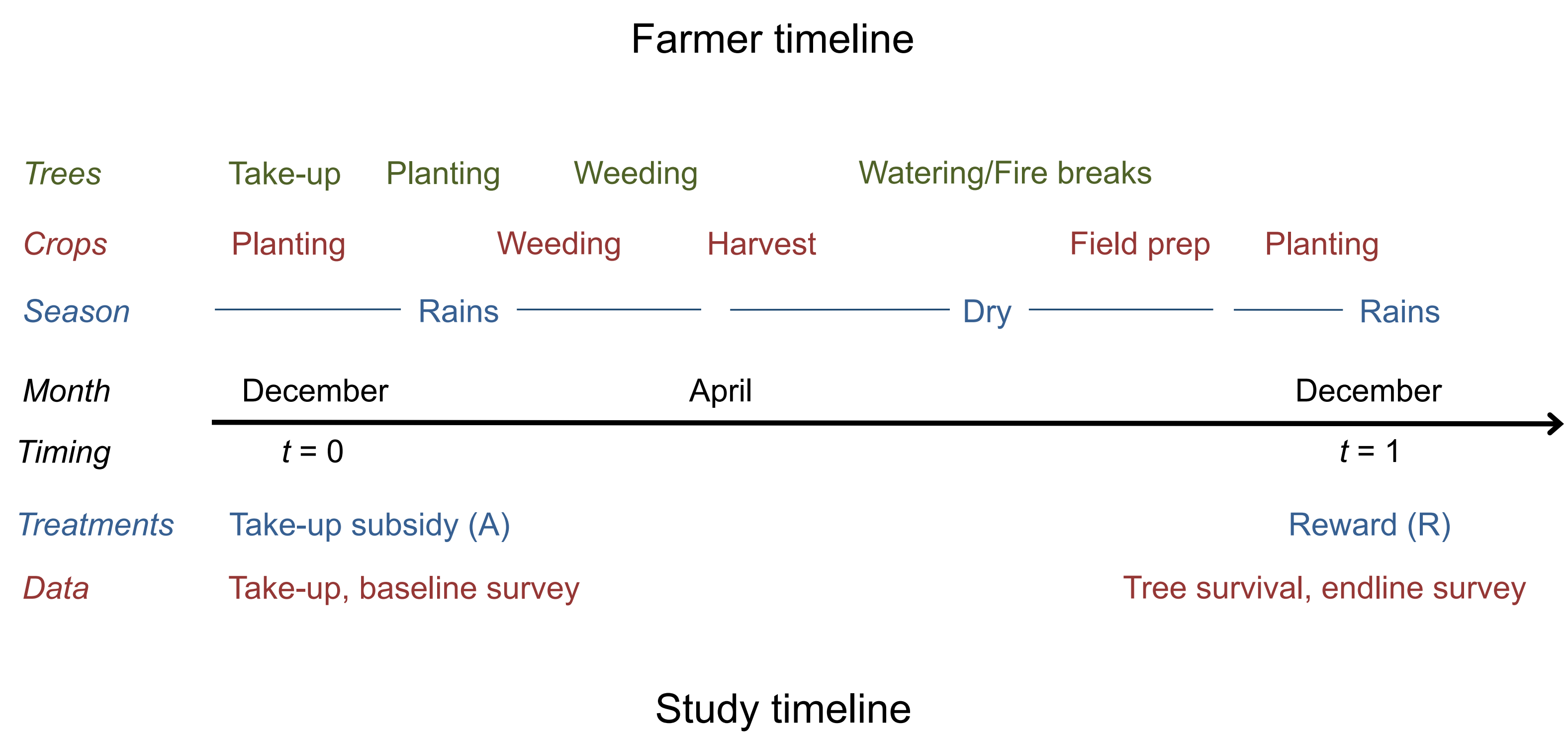
Table A.4.2: Incentive spillovers within group

Dependent variable is tree survival		
	(1)	(2)
Average reward in group (excl. own)	0.262	0.578*
	(0.240)	(0.317)
Own reward	0.319**	0.642**
	(0.0569)	(0.253)
Group reward x own reward		-0.0230
		(0.0182)
N	1088	1088

Notes: OLS regressions of tree survival on average draw in farmer group, conditional on take-up, and own draw. Standard errors are clustered at the group level. * p<0.10 ** p<0.05 *** p<0.01.

A.5 Additional tables and figures

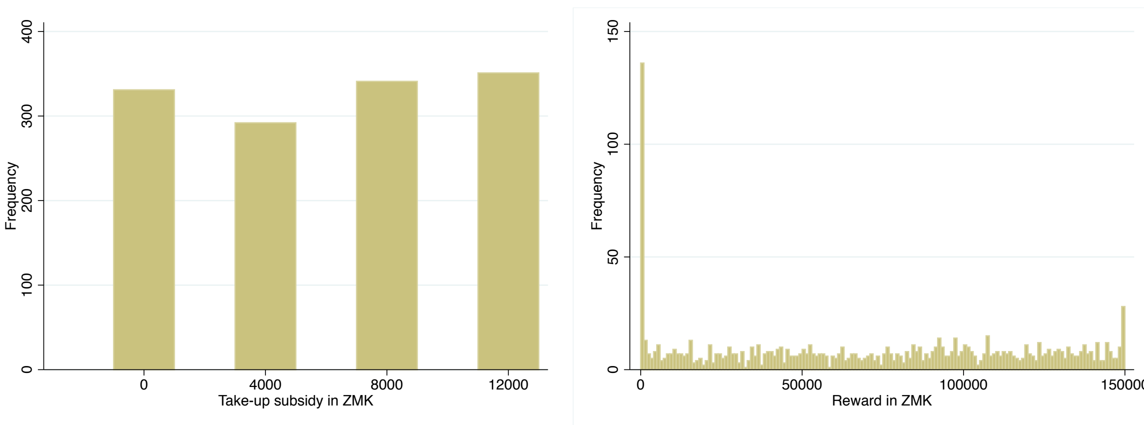
Figure A.5.1: Study timeline



Notes: Figure depicts the timing of (from top to bottom) activities undertaken by the farmer associated with tree cultivation, other crops, and the corresponding season and month. These are mapped on to the timing in the model and of the treatments and data collection.

Figure A.5.2: Distribution of treatments

(a) Take-up subsidy values (group-level) (b) Threshold reward values (individual-level)



Notes: Frequency of (a) take-up subsidy treatment values, randomized at the farmer group level in increments of 4
(b) threshold reward values, randomized at the individual level in increments of 1,000 ZMK. The threshold rewa
observations > 150,000 ZMK, which were awarded in error and are top-coded at 150,000 in the analysis. Twenty
draws were set at 0 ZMK.

Table A.5.1: Balance

	A=0 Mean [SD] (1)	A > 0 (2)	R=0 Mean [SD] (3)	Reward > 0 (4)	Surprise=0 Mean [SD] (5)	Surprise reward (6)	N (7)
Respondent is head of household	0.738	0.000	0.694	0.000	0.703	0.037	1291
	[0.441]	[0.003]	[0.463]	[0.0003]	[0.457]	[0.025]	
Age, respondent	37.853	-0.073	37.439	0.0071	37.238	1.296*	1265
	[13.733]	[0.100]	[12.808]	[0.0073]	[13.693]	[0.710]	
Female headed household	0.136	0.001	0.149	0.0001	0.119	0.009	1291
	[0.343]	[0.003]	[0.358]	[0.0002]	[0.324]	[0.017]	
Years of education, respondent	5.343	-0.039	5.284	0.0033	5.364	-0.01	1291
	[3.281]	[0.028]	[3.133]	[0.0022]	[3.334]	[0.178]	
Household size	5.46	-0.035**	5.328	0.0003	5.243	0.211*	1291
	[2.144]	[0.015]	[2.390]	[0.0013]	[2.218]	[0.114]	
Ordinal discount rate (1 - 5)	2.455	-0.001	2.538	0.0004	2.43	0.106	1261
	[1.629]	[0.013]	[1.714]	[0.0010]	[1.613]	[0.095]	
Non-agricultural assets	9.802	-0.101**	9.343	-0.0013	9.077	0.317	1291
	[5.797]	[0.041]	[5.625]	[0.0034]	[5.510]	[0.287]	
Years working with Dunavant	4.231	-0.033	3.776	-0.0015	3.866	0.067	1291
	[3.753]	[0.035]	[3.404]	[0.0022]	[3.499]	[0.214]	
Total landholdings (hectares)	3.022	-0.022	2.881	0.000	2.874	0.055	1289
	[2.252]	[0.023]	[2.188]	[0.0014]	[2.255]	[0.115]	
Number of fields	2.87	0.002	2.866	0.000	2.884	-0.052	1291
	[1.065]	[0.009]	[1.194]	[0.0008]	[1.122]	[0.070]	
Average distance from home to plots	20.565	-0.086	19.416	-0.0202	18.411	0.896	1291
	[24.226]	[0.236]	[22.436]	[0.0122]	[20.639]	[1.071]	
Poor soil fertility	0.108	-0.001	0.104	0.000	0.106	-0.029*	1291
	[0.311]	[0.002]	[0.307]	[0.0002]	[0.308]	[0.017]	
Regular interaction with lead farmer	0.413	0.004	0.448	-0.0006**	0.411	0.008	1289
	[0.493]	[0.004]	[0.499]	[0.0003]	[0.492]	[0.029]	
Affiliated with CFU or COMACO	0.037	0.002	0.037	0.000	0.042	0.003	1291
	[0.189]	[0.002]	[0.190]	[0.0001]	[0.202]	[0.012]	
Prior knowledge of Faidherbia	0.679	-0.002	0.664	0.000	0.639	0.027	1291
	[0.468]	[0.004]	[0.474]	[0.0003]	[0.481]	[0.025]	
Prior planting of Faidherbia	0.108	-0.001	0.09	0.0001	0.086	0.015	1291
	[0.311]	[0.002]	[0.287]	[0.0002]	[0.281]	[0.014]	
Knowledge of risks to tree survival	1.722	-0.005	1.701	0.000	1.649	-0.014	1291
	[0.906]	[0.006]	[0.785]	[0.0005]	[0.816]	[0.046]	
N	324	967	134	1041	613	678	

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. * p < 0.10 ** p < 0.05 *** p < 0.01.

Table A.5.2: Attrition across data collection phases

	Takeup Mean [SD]	Baseline	Endline	Tree monitoring
	(1)	(2)	(3)	(4)
Take-up subsidy	6.1674	0.0015	0.0000	0.0000
	[4.5383]	[0.0009]	[0.0020]	[0.0007]
Reward ('000 ZMK)	69.3347	0.0001	0.0000	0.0000
	[48.4713]	[0.0001]	[0.0001]	[0.0000]
Surprise reward treatment	0.5251	-0.0012	0.0093	0.0025
	[0.4996]	[0.0053]	[0.0122]	[0.0061]
N, outcome = 1	1314	1291	1232	1983

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 take-up (N=1091). For observations missing the reward variable (surprise reward treatment, no take up), a missing variable dummy for the reward is added to the regression. Reported coefficients are among non-missing reward values. * p < 0.10 ** p < 0.05 *** p < 0.01.

Table A.5.3: Farmer investments

	(1) Weeding	(2) Fire breaks	(3) Watering	(4) Burning
	Panel A: Outcome recorded during field visit			
Reward	0.0007*	0.0006	0.0008**	-0.0002
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Take-up subsidy	-0.0062	0.0071	0.0007	-0.0006
	(0.0055)	(0.0059)	(0.0050)	(0.0055)
Reward p-value	0.059	0.129	0.041	0.677
Mean	0.54	0.36	0.26	0.40
Observations	719	719	719	719
	Panel B: All contracted farmers			
Reward	0.0012***	0.0009***	0.0009***	0.0005*
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Take-up subsidy	-0.0025	0.0057	0.0009	0.0014
	(0.0045)	(0.0040)	(0.0035)	(0.0041)
Reward p-value	0.000	0.002	0.002	0.098
Dep. Var. Mean	0.36	0.24	0.17	0.27
Observations	1092	1092	1092	1092

Notes: OLS regressions of observed indicators of farmer effort on treatment variables. The reward and take-up subsidy are both measured in thousand ZMK. Regressions control for an effort monitoring indicator and cluster standard errors at the farmer group level. We omit other controls for comparability with Table 3. Results are similar if other controls are included. Note that these outcomes are recorded at the end of the contract period and do not necessarily reflect farmer investments at other points during the year. Panel A includes only non-missing observations. Farmers that reported zero surviving trees during the endline survey did not receive a field visit, and as a result, we do not observe whether they undertook these activities. Panel B replaces missing observations with zeros, which is valid if farmers with zero surviving trees were not undertaking these activities at the time of tree monitoring.

Table A.5.4: Correlation between farmer observables and program outcomes

	(1) Take-up	(2) Zero trees	(3) 35-tree threshold	(4) Tree survival
Household head at training	0.0640**	0.0178	0.0010	0.2770
	[0.0227]	[0.0366]	[0.0328]	[1.1939]
Female household head	0.0272	-0.0680	-0.0248	0.2885
	[0.0293]	[0.0415]	[0.0352]	[1.3798]
Respondent education	0.0002	0.0027	0.0079	0.3668*
	[0.0033]	[0.0049]	[0.0042]	[0.1744]
Household size	0.0105*	-0.0024	0.0063	0.1240
	[0.0048]	[0.0062]	[0.0057]	[0.1994]
Non-agricultural assets	0.0017	0.0033	-0.0003	-0.0606
	[0.0019]	[0.0029]	[0.0029]	[0.1018]
Years working with Dunavant	0.0061	0.0003	0.0077	0.1845
	[0.0034]	[0.0041]	[0.0041]	[0.1598]
Land size (hectares)	0.0051	-0.0037	-0.0043	-0.0261
	[0.0040]	[0.0065]	[0.0062]	[0.2424]
Number of fields	0.0057	-0.0287	-0.0042	0.6132
	[0.0088]	[0.0151]	[0.0126]	[0.5467]
Distance from home to plots	-0.0002	0.0004	0.0003	-0.0196
	[0.0006]	[0.0008]	[0.0007]	[0.0252]
Poor soil fertility	-0.0205	0.0680	-0.0237	-1.9494
	[0.0360]	[0.0568]	[0.0491]	[1.9100]
Sees YGL often	0.0261	-0.0620	-0.0142	0.9962
	[0.0181]	[0.0318]	[0.0270]	[1.0389]
Affiliated with CFU or COMACO	0.0039	-0.1479**	0.0635	4.0746
	[0.0427]	[0.0513]	[0.0709]	[2.3833]
Prior knowledge of Faidherbia	0.0356	-0.0429	0.0442	0.8228
	[0.0253]	[0.0349]	[0.0321]	[1.1978]
Prior planting of Faidherbia	-0.0603	-0.1309**	0.0779	5.0115*
	[0.0365]	[0.0435]	[0.0573]	[2.1682]
Knowledge of risks to tree survival	0.0152	-0.0319	0.0402**	1.7570**
	[0.0114]	[0.0174]	[0.0128]	[0.5794]
Constant	0.3462***	0.6850***	-0.0318	2.6194
	[0.0841]	[0.0901]	[0.0762]	[3.0433]
R squared	0.1879	0.0835	0.0750	0.1083
N	1288	1080	1080	1080
Dep. Var. Mean	0.8392	0.3574	0.2528	17.60

Notes: OLS regressions of outcomes on observables collected as part of the baseline survey, during training. The outcome in column 2 is an indicator for zero surviving trees and in column 3 is an indicator for reaching the reward threshold ($\geq$ 35 trees). All columns include controls for the experimental treatments: subsidy level, reward level, reward timing and monitoring. * p<0.10 ** p<0.05 *** p<0.01.

Table A.5.5: Learning

	Zero surviving trees		
	(1)	(2)	(3)
Knowledge var is:	Own msangu at baseline	Group msangu at baseline	Knows any risks to trees
Knowledge variable	-0.0600	0.0187	-0.1534***
	[0.0791]	[0.0714]	[0.0450]
Positive price (A<12,000)	0.0447	0.1079	-0.0252
	[0.0468]	[0.0747]	[0.0549]
Interaction	-0.1039	-0.1265	0.1134**
	[0.0944]	[0.0908]	[0.0554]
Observations	1080	1080	1080

Notes: OLS regressions of a binary indicator of whether the farmer had zero surviving trees, conditional on take-up. The regressions include a measure of baseline knowledge of the trees, an indicator for whether the take-up subsidy treatment resulted in a positive price for the seedlings and the interaction of the two. The regressions control for other variables shown in the balance table and for treatments. The knowledge variables are (1) whether the farmer had any msangu planted at baseline, (2) whether any farmer in the same farmer group had msangu planted at baseline and (3) whether the farmer could describe at least one risk to tree survival at baseline.

Table A.5.6: Procrastination

Dependent variable:	Take-up	Survival	Survival	Take-up	Take-up
	(1)	(2)	(3)	(4)	(5)
	Panel A: Self-described procrastinator				
Binary Procrastination Measure	-0.0091	-0.8574	-1.3361	-0.0305	-0.0016
	[0.0209]	[1.1563]	[1.1410]	[0.0442]	[0.0651]
Take-up subsidy	0.0219***	-0.0513	-0.0184	0.0202***	0.0233***
	[0.0044]	[0.2004]	[0.1952]	[0.0046]	[0.0058]
Reward in '000 ZMK			0.0669***		0.0007**
			[0.0114]		[0.0003]
Procrastination x subsidy				0.0034	-0.0028
				[0.0047]	[0.0069]
Constant	0.4524***	9.0534***	4.3205	0.4651***	0.3980***
	[0.0780]	[3.1850]	[3.1716]	[0.0801]	[0.1088]
N	1274	1071	1071	1274	603
	Panel B: Reports procrastination on other activities				
Binary Procrastination Measure	-0.0302	0.0622	0.1404	-0.0938	-0.0521
	[0.0278]	[1.2728]	[1.2767]	[0.0629]	[0.0767]
Take-up subsidy	0.0231***	-0.1045	-0.0745	0.0197***	0.0205***
	[0.0044]	[0.2021]	[0.1958]	[0.0043]	[0.0050]
Reward in '000 ZMK			0.0686***		0.0006*
			[0.0119]		[0.0003]
Procrastination x subsidy				0.0102	0.0076
				[0.0066]	[0.0080]
Constant	0.4589***	8.8880***	3.8534	0.4766***	0.4333***
	[0.0801]	[3.1334]	[3.1569]	[0.0806]	[0.1014]
N	1223	1030	1030	1223	576

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 *** p<0.01.

A.6 Appendix references

Ashraf, N., J. Berry, and J. M. Shapiro (2010). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review 100*(5), 2383–2413.

Benabou, R. and J. Tirole (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies 70*(3), 489–520.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Berry, J., G. Fischer, and R. P. Guiteras (2012). Eliciting and utilizing willingness to pay: Evidence from field trials in northern Ghana. *Working Paper*.

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics 125*(1), 1–45.

Milgrom, P. and J. Roberts (1986). Price and advertising signals of product quality. *The Journal of Political Economy*, 796–821.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge University Press.