CSAE WPS/2010-37

Heterogeneous returns and the persistence of agricultural technology adoption

November 2010

Andrew Zeitlin,[*†] Stefano Caria,[‡] Richman Dzene,[§]

Petr Janský,[**] Emmanuel Opoku[††], and Francis Teal[†]

Abstract

In this paper we explore whether low rates of sustained technology use can be explained by heterogeneity in returns to adoption. To do so we evaluate impacts of the Cocoa Abrabopa Association, which provided a package of fertilizer and other inputs on credit to cocoa farmers in Ghana. High estimated average productive impacts for treated farmers are found to be consistent with negative economic profits for a substantial proportion of the treated population. By constructing an individual-specific measure of returns, we demonstrate that low realized returns among adopters are associated with low retention rates, even after conditioning on output levels and successful repayment. The results are consistent with the hypothesis that high average returns mask substantial and persistent heterogeneity, and that farmers experiment in order to learn about their idiosyncratic returns. (JEL codes: O13, O33, Q12, Q16)

[*] Corresponding author. Email: andrew.zeitlin@economics.ox.ac.uk.
[†] Centre for the Study of African Economies, Department of Economics, University of Oxford
[‡] Overseas Development Institute
[§] Ghana Institute of Management and Public Administration
[**] Charles University in Prague
[††] Ghana Cocoa Board

1 Introduction

What constrains investment in agricultural technologies? This question is important – and stubbornly persistent – in development economics. Not only does agriculture continue to represent the primary source of income for many of the world's poor, but low adoption rates of agricultural technologies, such as fertilizers and improved seed varieties, have accompanied the stagnation of agricultural productivity in Africa in particular (World Bank 2008).

There is little dispute that there exist agricultural technologies with high expected returns in many Sub-Saharan contexts. This view is supported by a growing body of evidence. Notably, Esther Duflo, Michael Kremer and Jonathan Robinson report experimental evidence of a mean seasonal return of 36 percent to fertilizer use among maize farmers in the Busia District of Kenya (Duflo et al. 2008). And yet rates of fertilizer use are low: fewer than 24 percent of farmers in Duflo and coauthors' study had used fertilizer in the preceding year. Even where supposedly high-return technologies do get adopted, many farmers abandon them. In a distinct sample of Kenyan maize farmers, Tavneet Suri documents that 30 percent of farmers switch into and out of the use of hybrid seeds in a given year (Suri 2007). In Ethiopia, Stefan Dercon and Luc Christiansen find that, while only 22 percent of farmers use fertilizer in a given year, a further 14 percent of farmers in the final round of their survey were not using fertilizer in spite of having done so in previous survey rounds (Dercon & Christiaensen 2007). Low rates of adoption and lack of sustained use of the technology combined with high rates of return to those technologies therefore present a puzzle.

Several mechanisms have been put forward to explain observed patterns of agricultural technology adoption. Processes of social learning have been much studied (Bandiera & Rasul 2006, Conley & Udry 2010, Foster & Rosenzweig 1995, Munshi 2004). If social learning is sufficiently important, low-adoption equilibria may persist in spite of potentially high returns. Alternative theories include credit and supply-side constraints (Moser & Barrett 2006, Zerfu & Larson 2010). In Kenya, Duflo and co-authors find evidence consistent with the view that time inconsistency in farmers' preferences causes inefficient adoption decisions (Duflo et al. 2009).

While they do address important elements of observed adoption patterns, these theories are generally not well equipped to explain why adoption is not sustained. In the most common form of learning model, for example, farmers only adopt technologies when they know how to use them effectively, and this knowledge, once acquired, is never lost (Foster & Rosenzweig 1995, Jovanovic & Nyarko 1996). Likewise, instability in the supply of inputs alone seems an ad hoc explanation, and one incapable of explaining the widespread failure of farmers to persistently adopt profitable technologies even in cases where farmers have accessed them in the past. One exception is Dercon and Christiaensen (2007), who argue that year-to-year variation in the ability of households to bear risks associated with high-return technologies may explain instability in their use. Even so, if the typical farmer experiences such high returns as have been reported in the literature – a premise that we revisit in this paper – one would expect such technologies, once established, to pay for themselves. Indeed, these stylized facts lead

Duflo to assert that "prima facie, neither limited liability nor risk aversion seem capable of explaining such a low level of fertilizer use" (Duflo 2003).

In this paper, we test the hypothesis that persistent differences in returns across farmers play a key role in determining sustained adoption of new technologies. We do so using a unique dataset to study a non-profit initiative, the Cocoa Abrabopa Association, which alleviated credit constraints to the adoption of the *hi-tech* package of inputs (a specific combination of fertilizer, insecticide, and fungicide) among cocoa farmers in Ghana.

The role of treatment effect heterogeneity in explaining low adoption rates in African agriculture is understudied in the literature, with papers by Tavneet Suri (Suri 2007) and Duflo, Kremer and Robinson (Duflo et al. 2008) providing two notable exceptions. Using an observational, panel dataset of Kenyan maize farmers, Suri estimates a model that allows for heterogeneous returns to fertilizer. Suri's econometric method allows her to estimate a mean return to fertilizer use for four subgroups, which are defined by their adoption and disadoption histories in each of the four waves of her data. There is a non-monotonic relationship between the adoption rates and expected returns of these subgroups in her data. Suri makes sense of this by arguing that transaction costs are particularly high where returns are highest.

Rather than estimate mean returns for observable subgroups, we characterize the full distribution of returns for those who adopt. We do so in two ways: by estimating quantile treatment effects, and by constructing a proxy for experienced returns at the individual level. Quantile treatment effects and our panel-based proxy for individual returns provide a means of getting at this individual distribution under alternative, and non-overlapping, assumptions. We argue that a full characterization of the distribution of individual returns is important for understanding the relationship between returns adoption rates. High expected returns, even within an observable subgroup, may be driven by a few individuals with very high returns, and so may be consistent with negative returns to adoption for a majority of members in that subgroup. To understand the relationship between returns and adoption, we are interested in the fraction of farmers with positive (net) returns. Methods that return an average treatment effect – even within subgroups – do not yield this information.

Duflo, Kremer and Robinson (2008) also estimate quantile treatment effects. They find that net returns to fertilizer adoption are negative for 13.5 percent of farmers, in spite of the estimated 36 percent seasonal return.[1] Because their design is based on randomized allocation of fertilizer across the plots of a given farmer, Duflo et al. can estimate farmer-specific returns. Such an experimental approach provides an ideal source of variation in fertilizer use with which to identify internally valid estimates of the average returns to its adoption. However, randomized allocation of fertilizer does not by itself provide exogenous variation in the returns to fertilizer experienced by individual farmers. The key

[1] As pointed out by Andrew Foster and Rosenzweig, these "net" returns do not account for labor costs (Foster & Rosenzweig 2010). If labor or other complementary inputs also increased in response to fertilizer use, then they will overestimate the true net returns, and underestimate the fraction of farmers for whom net returns are negative.

contribution of this paper is to demonstrate a relationship between the returns experienced by individual farmers and their subsequent decisions to sustain use of the technology.

We take as our starting point that the literature is far from conclusive on three questions: (1) How heterogeneous are the rates of return to agricultural technologies, such as fertilizer? (2) Does heterogeneity in returns affect the sustained adoption of such technologies, beyond farmers' initial experimentation? And (3) is this heterogeneity in realized returns caused by transient shocks, or does it reflect persistent differences in the suitability of a technology across farms and farmers? In the case of Ghanaian cocoa, we answer each of these in the affirmative: heterogeneity in the returns to fertilizer use is substantial, it affects continued adoption, and it reflects persistent differences across farmers.

We proceed as follows. Section 2 describes the intervention, the data and the quasi-experimental setting that will be used to estimate impacts. Section 3 presents our estimates of the average treatment effect, demonstrating robustness to a range of identifying assumptions. Having established large positive returns on average, we turn in Section 4 to demonstrating the heterogeneity of these treatment effects. In Section 5, we test for a relationship between experienced treatment effects and program retention. Section 6 concludes.

2 Context, data and quasi-experimental design

In 2006, the Cocoa Abrabopa Association (CAA), a not-for-profit subsidiary of Wienco Ghana Ltd, began a program of distributing inputs on seasonal credit to cocoa farmers in Ghana. With the support of the Ghana Cocoa Board, CAA provided farmers with access to two acres' worth of a package of fertilizer, pesticides, and fungicides. This specific bundle of inputs, known as the *hi-tech* package, had been promoted by the Cocoa Research Institute of Ghana since 2001, though problems of poor repayment rates had limited distribution.[2] CAA provided these inputs to groups of between 8 and 15 farmers on a joint liability basis, with dynamic incentives: groups that failed to repay in full would be suspended for a minimum of one year, while those that repaid successfully would be given four acres' worth of inputs in the following year, subject to approval of a CAA field officer. In addition to these physical inputs, farmers in the first year of membership were advised on their proper application by a CAA promoter, and some business training would be provided by Technoserve Ghana. Judged by its expanding membership rolls, the program has been wildly successful: from 1,440 farmers in 2006, CAA expanded to a membership of 18,000 farmers by 2009 (Cocoa Abrabopa Association 2009).

To identify the impact of CAA membership on farmer incomes, we took advantage of the fact that much of CAA's expansion during this time was at an extensive margin: it involved expansion into new villages. CAA's expansion operates on an annual cycle, as follows. Promoters first arrive in a new village in January of a given year, and by February farmers make their decisions to opt into the program (or not),

[2] Although the use of these broad categories of inputs is not new to Ghanaian cocoa farmers, the particular configuration was. Evidence from other contexts (Duflo et al. 2008) shows that economic returns can be highly sensitive to precise quantities and combinations of inputs used.

forming groups accordingly. Inputs begin to arrive in May, but the harvest does not take place until October, with repayment of inputs due by December of that year.

We visited farmers in September of 2008 and 2009 to observe the outcomes for the 2007/08 and 2008/09 seasons, respectively. In each wave of the survey, we conducted a representative sample of two types of villages: those that had been reached by CAA for the first time in the prior year, and those that had been reached by CAA for the first time in the current year, i.e., the growing season that was ongoing at the time of our visit.[3] The former had experienced one full season since the arrival of Abrabopa and had made their membership decisions for the following season, but had not yet harvested any cocoa in the second year of exposure by the time of our visit. The latter had made membership decisions for the season in progress at the time of our survey, but had not yet experienced the results of those decisions.

In each type of village, we conducted representative samples of two populations of farmers: those who joined in the first year of its availability in their village, whom we call *early adopters*, and those who opted not to join CAA in the same year, whom we call *early non-adopters*. The resulting sample used in estimation is given in Table 1.

Table 1. Estimating sample, by survey round and membership classification

Year village first visited by Abrabopa	Farmer's adoption decision in year of first visit	Number of observations, by survey round 2007/08 season	2008/09 season
2007	Adopt	82	0
	Do not adopt	41	0
2008	Adopt	88	72
	Do not adopt	42	29
2009	Adopt	0	95
	Do not adopt	0	37

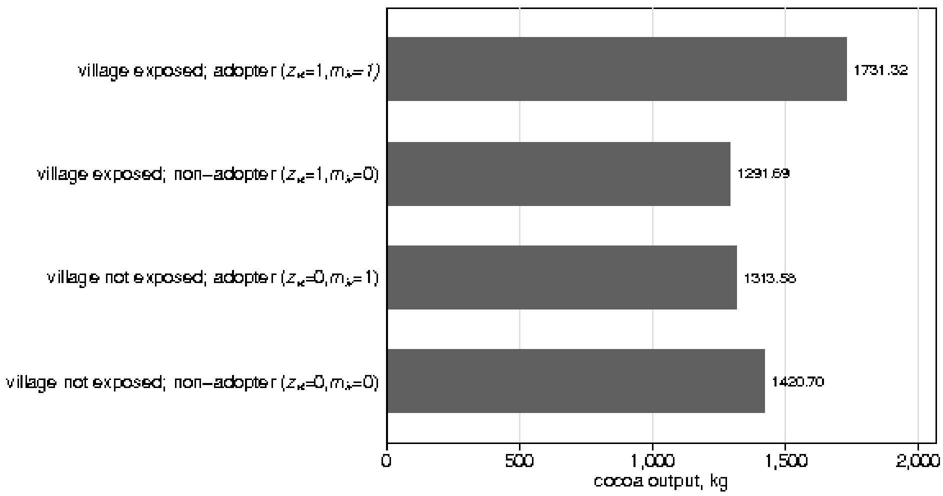
Note: survey round refers to the most recent completed harvest as of the time of each survey. 2007 denotes the 2007/08 cocoa season; 2008 denotes the 2008/09 cocoa season.

As will be described in detail in the next section, we use a cross-sectional difference-in-differences approach to estimate the average effect of the first wave of Abrabopa membership on early adopters. To do so, we pool data from both survey rounds. We then define two key variable. First, we define a dummy variable, z_{vt}, to indicate exposure to Abrabopa in village v and year t. Given the sampling strategy in Table 1, any village in our sample with $z_{vt} = 0$ must have $z_{v,t+1} = 1$. Second, we define m_{ivt} as an indicator of the adoption decision of farmer i in village v and year t. To denote early adopters, who adopt in the first year of exposure, we drop the time subscript and define $m_{iv} = \mathbf{1}\{m_{ivt} = 1, z_{vt} = 1, z_{vs} = 0\}$ for some t and for all $s < t$. Early adopters choose to adopt in the first year of availability in their village, but not necessarily in subsequent years, as we shall see. In the sample used to estimate impacts on farmer production – which restricts attention to the year of exposure or the year

[3] In the 2009 survey, we also revisited farmers from the previous survey who had by then been exposed to the program for two seasons. To focus on a comparable set of "early adopters", and in light of the possible cumulative effects of sustained fertilizer use, we do not make use of these observations in this paper. See Opoku and coauthors (Opoku et al. 2009) for further details of the survey.

immediately prior – treatment w_{ivt} occurs if both the individual is an early adopter and Abrabopa is present in the village: $w_{ivt} = m_{iv} z_{vt}$.

Figure 1. Average cocoa output, by treatment status in year of survey



The basis for our difference-in-differences identification strategy is illustrated in Figure 1, which shows our four categories of farmers. The first $(z_{ivt} = 1, m_{iv} = 1)$ are those who receive treatment (i.e., they have adopted the *hi-tech* package). The second $(z_{ivt} = 1, m_{iv} = 0)$ were offered the treatment but did not adopt. The third $(z_{ivt} = 0, m_{iv} = 1)$ are farmers who will adopt in the first year that Abrabopa reaches their village, but whose villages have not completed a season under treatment at the time of the survey round under consideration. And fourth are those farmers who will choose not to join in the first year of exposure, and whose villages have not yet been exposed to Abrabopa $(z_{ivt} = 0, m_{iv} = 0)$. Given this setup, we can read off a difference-in-differences estimate of program impacts directly from Figure 1. We estimate the average treatment effect on the treated as

$$\begin{aligned} \tau_{ATT} = \quad & E(y_{ivt}|m_{iv} = 1, z_{vt} = 1) - E(y_{ivt}|m_{iv} = 0, z_{vt} = 1) \\ & - [E(y_{ivt}|m_{iv} = 1, z_{vt} = 0) - E(y_{ivt}|m_{iv} = 0, z_{vt} = 0)]. \end{aligned} \tag{2.1}$$

The part of (2.1) gives the differences in output between those who were early adopters and those who were not, in villages exposed to Abrabopa. This within-village difference nets out one potential form of selection bias, arising from any common, village-level productivity variables that may be correlated with the timing of Abrabopa exposure. Still, this may be a biased estimate of the treatment effect if relatively productive farmers were more likely to become members. By subtracting the second part of (2.1), the pre-treatment difference between early adopters and non-adopters in villages not yet exposed to Abrabopa, we can account for this second form of selection bias, arising from within-village selection.

In the pooled cross-section without any controls, the average treatment effect of Abrabopa adoption in equation (2.1) can be read off from Figure 1 as 546.75 kg.[4] From the figure, neither form of selection bias appears very strong: in villages not yet experiencing output under treatment, those who go on to become early adopters are, if anything, worse off than no-adopters. Moreover, non-adopters in villages under treatment are, if anything, worse off than non-adopters in villages not yet treated.

Sampled farmers also provided information on a range of socio-economic characteristics and agricultural practices. These data are summarized in Appendix Table A1, where we present summary statistics by survey round and treatment status. Table A1 includes the prima facie evidence of Abrabopa's impacts that underlie Figure 1: early adopters' output exceeds that of non-adopters in exposed villages in both survey waves, while early adopters' output levels in the season before their village is exposed to Abrabopa are higher than non-adopters in 2008 (1615 vs 1460 kg) but lower in 2009 (1034 vs 1376 kg). Rates of fertilizer use are low (less than 50 percent) among farmers not yet reached by Abrabopa. Farms in this sector are typically small: farmers have an average of approximately four hectares of land devoted to cocoa trees. Education levels are low, just under 40 percent of farmers having completed education beyond primary level, and just over 20 percent of farmers in the sample are female. Given the quasi-experimental setting, we defer a discussion of covariate balance across treatment and control groups to the following section, where we present a more detailed explanation of our identification strategy.

3 Average returns

To make clear the identifying assumptions underlying our estimates of average returns, consider the following model for potential outcomes of gross output under two counterfactual scenarios – with and without the *hi-tech* inputs ($w = 1,0$ respectively):

$$y_{0ivt} = \mu_0 + \eta_i + \lambda_{vt} + \varepsilon_{0ivt} \tag{3.1}$$

$$y_{1ivt} = \mu_1 + \eta_i + \lambda_{vt} + \varepsilon_{1ivt} \tag{3.2}$$

for farmer i in village v and year t. For the time being, we ignore the role of observed, farmer-specific covariates. The η_i give farmer-specific, time-invariant unobserved characteristics, while the λ_{vt} capture village-year unobserved shocks to productivity; we will be concerned about the potential correlation of both with treatment status.[5] Without further loss of generality, we assume that the $E[\eta_i] = E[\lambda_{vt}] = E[\varepsilon_{wivt}] = 0$, for $w = 0,1$, so that the difference $\tau_{ATE} = \mu_1 - \mu_0$ gives the average treatment effect in this population. The quantity $\tau_{ATT} = \mu_1 - \mu_0 + E[\varepsilon_{1ivt} - \varepsilon_{0ivt} | w_{ivt} = 1]$ gives the average treatment effect on the treated.

---

[4] This differs from the estimate in column (2) of **Table 2** only because the estimates in that table include a control for the year of survey.

[5] For counterfactual states, $w = 0,1$, we can decompose the state-specific error term, ε_{wivt}, into two parts, $\alpha_{wiv} + u_{wivt}$, where the first of these represents a time-invariant, individual-specific return ("essential heterogeneity"), which is potentially knowable by the farmer. The second component, u_{wivt}, is time-varying and captures a source of risk. Both can impact future technology decisions: the former affects subjective perceived returns, while the latter affects farmers' liquidity, buffer stocks, etc.

Observed outcomes are therefore given by the switching regression, $y_{ivt} = y_{0ivt} + (y_{1ivt} - y_{0ivt})w_{ivt}$. Substituting in equations (3.1) and (3.2) yields

$$y_{ivt} = (\mu_1 - \mu_0 + \varepsilon_{1ivt} - \varepsilon_{0ivt})w_{ivt} + \mu_0 + \eta_i + \lambda_{vt} + \varepsilon_{0ivt}. \qquad (3.3)$$

Examination of equation (3.3) clarifies the nature of the selection problem that must be addressed in estimating the ATT. A regression of y_{ivt} on w_{ivt} returns a consistent estimate of τ_{ATT} only if $[\eta_i + \lambda_{vt} + \varepsilon_{0ivt}|w_{ivt}] = E[\eta_i + \lambda_{vt} + \varepsilon_{0ivt}] = 0$. Thus consistent estimates of τ_{ATT} are possible even if individual-specific returns to adoption $(\varepsilon_{1ivt} - \varepsilon_{0ivt})$ are correlated with adoption choices, w_{ivt}. But the assumption required for identification of the ATT from a regression of output on treatment alone will fail if adoption is correlated *either* with village-level differences in productivity or with the idiosyncratic productivity of farmers.

To see how we address these sources of selection bias in practice, we first formalize the process by which membership is determined. Recall that we define treatment, w_{ivt}, as the effect of the first year of use of the hi-tech package, in light of the potential for accumulation of impacts over years. In our data, we restrict attention to village years, vt, in which either (a) Abrabopa has never had any members in that village before, and Abrabopa will have its first members in village v in year $t + 1$; or (b) Abrabopa has its first ever members in village v in year t. Consequently we examine first-year impacts only on the subpopulation of individuals who join Abrabopa in the first year that it is available in their village.

For this subpopulation, use of Abrabopa's hi-tech inputs is the product of two factors: firstly, that Abrabopa visits the individual's village, v, in year t, and secondly, that the individual joins in that year. Let z_{vt} be a dummy variable indicating the presence of Abrabopa in village v in year t, and let m_{iv} be a dummy variable indicating that individual i in village v is the "type" who joins Abrabopa in the first year in which it is available in their village. Thus $w_{ivt} = z_{vt}m_{iv}$.

Our identification strategy rests on two key features of our data. The first of these is the ability to observe the *future* membership decisions of individuals in villages that have not yet been visited by Abrabopa at the time of the output realization y_{ivt}.[6] The second of these is the ability to observe the productive outcomes for a representative sample of those who do *not* join Abrabopa in any given village-year.

We use the first of these features to address potential correlation between the individual-specific unobservables and treatment status, arising through individual selection into Abrabopa. The second allows us to address the potential correlation between village-level characteristics and treatment status, arising through the non-random roll-out of Abrabopa coverage.

To do so, we assume that the process by which farmers are selected into membership is constant over time, with respect to unobserved characteristics that differentiate them from village-mean productivity:

$$E[\eta_i + \varepsilon_{0ivt}|m_{iv} = 1, z_{vt} = 1] = E[\eta_i + \varepsilon_{0ivt}|m_{iv} = 1] \qquad (A1)$$

[6] For a similar use of future adoption decisions to address selection problems in a pipeline evaluation, see, e.g., Erica Field (Field 2005).

If those who join Abrabopa when it first reaches their villages in year t are the best farmers in those villages, then farmers who join Abrabopa upon its arrival in their villages in year $t+1$ are also the best farmers in those villages.

If we were also willing to assume that the roll-out of Abrabopa availability was as good as random with regard to village productivity levels ($E[\lambda_{vt}|z_{vt}] = 0$), then a comparison of current and future program members would suffice to identify the ATT. We operationalize this as a first, heuristic identification strategy by restricting attention to current and future members, and regressing output on treatment status.

Effective randomness of roll-out is a strong assumption, however. We are able to relax this by making use of data on non-members in program villages. In essence, we can use mean outcomes of those who do not join Abrabopa in current and future program villages to estimate the village-specific effect. Our estimates under this identification strategy are then a form of *difference-in-difference* estimates: we compare within-village differences between those who join Abrabopa and those who do not, in villages that have just been reached by Abrabopa in year t and those that will only be reached by Abrabopa in year $t+1$.

This strategy requires an auxiliary assumption that there are no externalities from the presence of the program on non-members – akin to the standard *stable unit treatment value* assumption. Formally, we require that

$$E[\eta_i + \varepsilon_{0ivt}|m_{iv} = 0, z_{vt} = 1] = E[\eta_i + \varepsilon_{0ivt}|m_{iv} = 0, z_{vt} = 0]. \qquad \text{(A2)}$$

Since we are interested here only in impacts in the first year of the program's presence in a village, we believe this to be a plausible assumption. Neighboring farmers will not have had an opportunity to observe program impacts and make any corresponding adjustments in their own production.[7]

Given assumptions (A1) and (A2), we operationalize our difference-in-differences estimator in two ways. Most directly, we regress output on dummy variables for adoption in the first year of exposure, m_{iv}, current exposure to Abrabopa, z_{vt}, and their interaction, $w_{ivt} = m_{iv}z_{vt}$, which is synonymous with treatment in our estimating sample. Under assumptions (A1) and (A2), a regression of the form

$$y_{ivt} = \beta_0 + \beta_z z_{vt} + \beta_m m_{iv} + \beta_w m_{iv} z_{vt} + e_{ivt} \qquad (3.4)$$

consistently estimates the ATT as the coefficient β_w. An alternative is to use a fixed effects estimator for equation (3.4), in which case the village-level exposure variable, z_{vt}, is not identified, but the treatment effect for early adopters remains identified by the interaction of indicators for the early adoption decision of the individual and village-level exposure. The latter may improve precision in a context where there is substantial heterogeneity in production across villages even within the set of exposed or unexposed villages.

[7] Conley and Udry demonstrate that farmers do adapt their technology choices in response to 'news' about their neighbors' levels of production (Conley & Udry 2010). As in their estimation strategy, we rely on farmers' inability to react until after output realizations have occurred.

Table 2. Estimates of average impact of CAA membership on members in the first year

VARIABLES	(1) Early adopters OLS	(2) Full sample DD	(3) Full sample FE	(4) Full sample, controls FE
treated: early adopter * current Abrabopa village, $m_{iv} * z_{vt}$	395.9*	547.6**	661.7***	527.4***
	(217.3)	(226.3)	(204.4)	(155.8)
early adopter, m_{iv}		-89.7	-222.8	-57.9
		(232.3)	(160.4)	(145.5)
current Abrabopa village, z_{vt}		-147.6		
		(293.2)		
constant	1533***	1582***	1405***	319.9
	(320.5)	(235.0)	(79.16)	(732.2)
Observations	337	486	486	359
R-squared	0.031	0.025	0.012	0.341
Number of village-years			32	29

Dependent variable is cocoa output, in kg. Robust standard errors in parentheses, clustered at village-level. Estimates in columns (1) and (2) contain dummy for survey round (not shown). Columns (3) and (4) include village-year fixed effects. Estimates in column (4) contain quadratic functions of farm size and farmer age, controls for household size, farmer education (JSS or greater), and farmer gender. Sample in column (1) is current and future members only; subsequent columns include those who do not join at any point. *** p<0.01, ** p<0.05, * p<0.1

Results of these estimates are presented in Table 2. In column (1), we estimate program impacts simply by regressing output levels on a dummy for current membership, while restricting the sample to current and future members of Abrabopa. As discussed above, this provides a valid estimate of the treatment effect only under the strong assumption that the roll-out is uncorrelated with unobserved location characteristics. The assumption of random roll-out yields an estimated average treatment effect on the treated of 359.9 kg.

In column (2) we relax this assumption by using data on both those who join Abrabopa when it first reaches their village and those who choose not to join in this initial year. This is the (cross sectional) difference-in-differences estimator described by equation (3.3) above, and returns consistent estimates of the treatment effect under assumptions (A1) and (A2). Non-members in villages first visited by Abrabopa in year $t + 1$ represent the omitted category. Several features are notable. We estimate an ATT of 547.6 kilograms, and this estimate is statistically distinguishable from zero at the 5 percent confidence level. We find only limited evidence for the two forms of selection bias considered. Non-adopters in villages already reached by Abrabopa, who have $z_{vt} = 1$, have output levels 147.6 kilograms less than non-adopters in villages not yet reached by Abrabopa; this is our estimate of $E[\lambda_{vt}|z_{vt} = 1]$. While this estimate is relatively imprecise, it does not provide support for the view that Abrabopa was systematically rolled out to more productive villages early on, or that non-adopters in exposed villages have benefited indirectly from the participation of their peers. Moreover, those who go on to be early adopters are, if anything, *worse* off than those who do not adopt at the first opportunity, in villages prior to exposure. Our estimate of $E[\eta_i|m_{iv} = 1]$ is -89.7 based on the identification strategy in column (2).

Estimates from the fixed effects specification are presented in column (3), where fixed effects are included for each village-year in our data to allow for the fact that some villages are observed both before and after exposure. This specification returns a similar estimate of the ATT, 661.7 kg, which is now significant at the 1 percent level. To put the magnitude of this effect in perspective, future members of Abrabopa have an average output of 1,464 kg: the estimated effect is equal to an approximately 45 percent increase in total production. With mean yields of 237 kg/acre among future joiners of the program, under the assumption that all impacts are concentrated on the 2 acres to which the inputs are supposed to be applied, this would represent an increase in yields on this land of nearly 140 percent.

In column (4) of Table 2, we augment our village-year fixed effects specification by including a vector of controls for observable farmer characteristics. We do not control for productive inputs, on the grounds that changes in labor and non-labor inputs mediate the causal effect of program membership on production (Foster & Rosenzweig 2010). More generally, controlling for post-treatment variables may introduce biases into estimates of causal effects (Heckman & Navarro-Lozano 2004, Pearl 2009). We include controls for farmer gender, an indicator variable for whether the farmer has attained junior secondary or higher education, and household size, as well as (quadratic functions of) farmer age and cultivated farm size. Given that it typically takes three years for cocoa trees to reach bearing age, we consider it reasonable to take our measure of cultivated farm size, which explicitly excludes trees too young to bear cocoa, as exogenous in this context. The resulting estimated treatment effect is reduced to 527.4 kg, but is still statistically significant in its difference from zero at the one percent confidence level. The similarity of these results to those from the village fixed effects specification suggest that unconfoundedness, conditional only on the village fixed effect and the first-year membership decision, is not an unreasonable assumption in this context. Relative to a mean output level of 1313 kg among eventual Abrabopa members in the year prior to exposure, this represents an increase in gross production of approximately 40 percent.

We can use a similar identification strategy to test whether participation in Abrabopa increases use of complementary inputs. If such impacts were to be found, then these should be taken into account in calculating returns to participation and technology use. For any input, x_{ivt}, used by farmer i in village v and year t, the effect of Abrabopa membership on input use can be implemented, as in column (3) of Table 2 and analogous assumptions to (A1) and (A2), by regressing x_{ivt} on early adoption, m_{iv}, and its interaction with current exposure, $w_{ivt} = m_{iv}z_{vt}$, in a specification with village-year fixed effects. We do so in particular for three types of labor input: household labor, hired labor, and *nnoboa* labor, a traditional labor-sharing arrangement. These results are reported in columns (1) to (3) of Table 3. In each case, point estimates are economically small, and we are not able to reject the hypothesis that these complementary input levels remain constant under participation in the program. In the absence of any such indirect costs, an impact of 527.4 kg corresponds to an economic return of 129 percent on the value of the inputs received.

Table 3. Intermediate inputs and covariate balance

VARIABLES	(1) Household labor days	(2) Paid labor days	(3) *Nnoboa* labor days	(4) Cocoa farm size, ha	(5) Age	(6) Female	(7) Household size	(8) Completed JSS or higher education
early adopter * current Abrabopa village, $m_{iv} * z_{vt}$	3.029	15.64	0.561	-0.365	1.594	-0.0707	0.178	0.0213
	(17.93)	(11.64)	(1.198)	(0.742)	(2.163)	(0.0728)	(0.262)	(0.0785)
early adopter, m_{iv}	-15.23	-0.420	1.197***	-0.805	2.418	-0.0330	0.0207	0.0594
	(11.26)	(7.481)	(0.425)	(0.535)	(2.469)	(0.0683)	(0.232)	(0.0941)
Constant	84.19***	26.36***	1.711***	4.657***	46.31***	0.269***	3.145***	0.525***
	(7.555)	(3.788)	(0.402)	(0.385)	(1.596)	(0.0453)	(0.148)	(0.0534)
Observations	482	482	482	510	368	368	368	368
R-squared	0.005	0.009	0.010	0.012	0.014	0.008	0.001	0.005
Number of village-years	32	32	32	32	29	29	29	29

Cluster-robust standard errors in parentheses, clustered at village level. All specifications include village-year fixed effects. Farm size defined as area under cocoa cultivation. *** p<0.01, ** p<0.05, * p<0.1

Finally, the same approach can be used to estimate pseudo-treatment effects on variables that should *not* respond to treatment (in the short term), such as farmer gender, age, and education, as well as cocoa farm size.[8] This is analogous to the standard practice of checking for 'successful' randomization in prospective experimental designs. Rejection of the null hypothesis of no pseudo-effect on any of these variables would suggest, minimally, the importance of controlling for them in estimating causal effects, as in column (4) of Table 2. Such a rejection might also give cause for concern about the assumptions underlying the identification strategy more generally (Heckman & Hotz 1989). However, as reported in columns (4) – (8) of Table 3, we are able to accept the null hypothesis of zero pseudo-effects on each of these predetermined variables.

4 Quantile treatment effects

Estimated returns to the CAA technological package are high among adopters, and yet 32 percent of first-year members in the sample were no longer members one year later. This appears to be a special case of the broader puzzle in agricultural technology adoption: if returns to adoption of this technology are indeed so high, and if credit constraints are not binding in this case, then why do so many members drop out?

Persistent heterogeneity in returns across individuals provides one explanation. If expected returns vary substantially across individuals – and if individuals are able to learn about their idiosyncratic returns – then low adoption could be an outcome of rational choices. The average treatment effect on the treated estimated above is not informative about the fraction of farmers expecting positive returns –

[8] Cocoa trees generally require a minimum of three years to reach bearing age.

what Heckman calls the "voting criterion" (Heckman 2010). In this section, we examine the extent to which high average treatment effects mask substantial heterogeneity in the ex-post distribution of impacts.

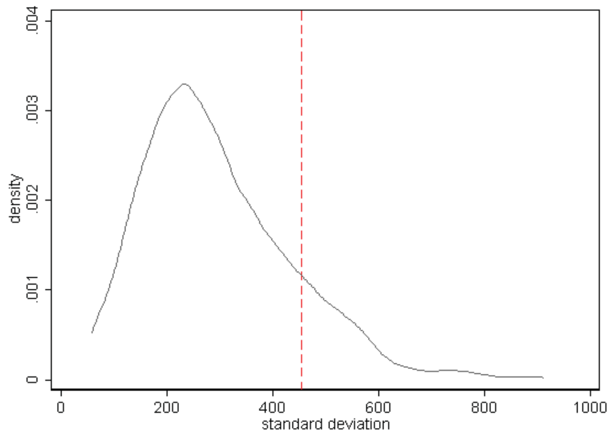
To do so, we conduct two exercises. First, following Heckman and co-authors (Heckman et al. 1997), we estimate the Hoeffding and Frêchet bounds on the amount of heterogeneity consistent with the marginal distributions observed in the data (Frêchet 1951, Hoeffding 1940). This allows us to obtain a lower bound on the variance of the treatment effect from the marginal distributions with and without treatment. The estimated lower bound on the amount of treatment effect heterogeneity is substantial in economic terms, though we are unable to reject a constant effects model given small sample limitations. These bounds exhaust the available information in the absence of further assumptions. Second, we consider the implications of assuming perfect positive dependence (PPD) in the outcome distributions for the quantiles of the impact distribution. Under this assumption, we show that in spite of the high average gross impacts, returns net of input costs are not significantly positive for a substantial proportion of those who join Abrabopa.

We begin by estimating Hoeffding and Frêchet bounds to test whether the estimated impacts in our data are consistent with a constant-effects model. Since Heckman and co-authors argued on behalf of doing so (Heckman et al. 1997), a handful of studies have applied this approach to estimate the distribution of treatment effects. In a developed-country context, Bitler and co-authors study the distribution of the effects of welfare reform on the unemployed (Bitler et al. 2006). Two recent articles have applied this approach in development economics, both studying the distribution of the effect of conditional cash transfers: Djebbari and Smith, on PROGRESA, and Dammert, on the Red de Protección Social in Nicaragua (Dammert 2009, Djebbari & Smith 2008).

We adapt the procedure outlined by Abbring and Heckman (Abbring & Heckman 2007) and as applied by Djebbari and Smith (Djebbari & Smith 2008) as follows. First, we obtain an estimate of the minimum amount of heterogeneity in impacts that is consistent with our data. We construct this measure as the standard deviation of the quantile treatment effects,[9] which is equal to 454 kg in our sample. Second, to test whether the observed impact standard deviation is consistent with a constant-effects model, we simulate the distribution of this statistic under the constant effects assumption. This is done, as in Djebbari and Smith, by keeping only control villages. For each simulation, we randomly assign half of the control villages to our synthetic treatment group; we use the estimated treatment effect from Table 2, column (2) to simulate treatment for the eventual Abrabopa members in those villages; and we repeat our cross-sectional, difference-in-differences estimation procedure to obtain an impact standard deviation on this synthetic sample.

[9] As Heckmen and co-authors argue (Abbring & Heckman 2007, Heckman et al. 1997), this is equal to the standard deviation of the treatment effects, $\Delta_{ivt} \equiv Y_{1,ivt} - Y_{0,ivt}$, under the assumption of perfect positive dependence, discussed below. PPD is the assumption about the joint distribution of (Y_1, Y_0) that minimizes the impact standard deviation. We base this procedure on estimates of the treatment effect at each decile, and do not use a finer division of the outcome distribution because of our small sample size.

Figure 2. Simulated distribution of estimated impact standard deviation under the constant effects model



Notes: Figure illustrates kernel density estimates of the simulated distribution of estimated standard impact standard deviation. Simulation is conducted under the null of a constant effects model, with average treatment effect as reported in Table 2. Results are for 400 repetitions, with 10 quantiles of the treatment distribution estimated in each repetition. Dotted line shows the value of the impact standard deviation in the actual data.

The results are illustrated in Figure 2. The median impact standard deviation across 400 simulations of the constant effects model is 337 kg, substantially smaller than that found in our data. However, in 24 percent of our simulated treatment effects, the constant effects model yields an impact standard deviation greater 454 kg, the value estimated from our actual data. Because our small sample does not estimate quantile treatment effects very precisely, we are unable to reject the constant effects model at conventional significance levels.[10]

Although we are unable to statistically reject the constant-effects model, we can use a quantile treatment effects approach to show that the heterogeneity in returns we do observe has economic implications. Differences in quantiles of the distribution of outcomes under treatment and control can be interpreted as quantiles of the treatment effect only under the assumption of perfect positive dependence. In this case, the treatment does not change the *ranking* of outcomes, so that the first quantile of the distribution without treatment, Y_0, represents the counterfactual for individuals in the first quantile of the distribution with treatment, Y_1. With treatment and control groups of different sizes, we compare impacts across quantiles of the outcome distribution, rather than directly matching individuals.

To do so, we continue with the identification strategy of Section 3. In particular, we estimate quantile treatment effects using a model that includes a dummy variable for presence of Abrabopa in the village

---

[10] To improve the power of this test, we have also simulated constant-effect treatment effects for the whole sample, including both the treated and the untreated, by subtracting the ATE from the realized output of the treated and then re-assigning all villages to treatment or control status and simulating treatment for members accordingly. The resulting increase in the sample size employed in each simulation reduces the fraction of simulated repetitions with an impact standard deviation greater than 454 kg to 0.138, and reduces the median estimated impact standard deviation to 266 kg.

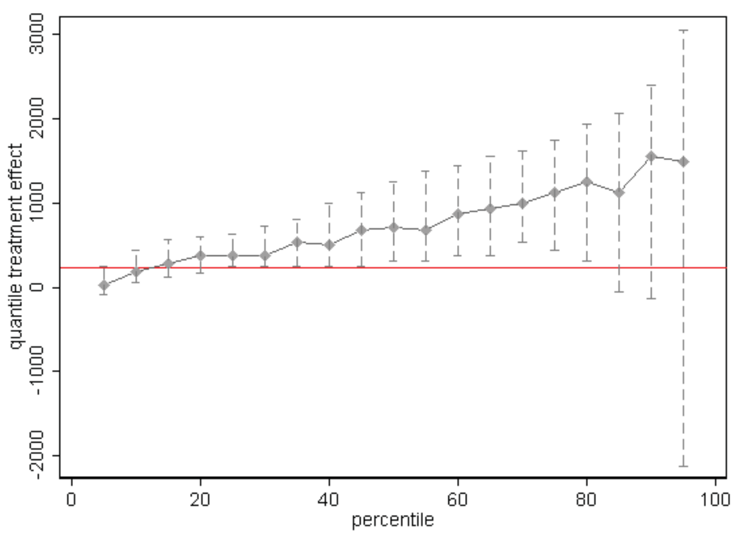
in the year under study, z_{vt}, as well as a dummy identifying those individuals who join Abrabopa in the first year of exposure in their village, m_{iv}. Treatment is denoted by the interaction of early adopters with village-level exposure: $w_{ivt} = m_{iv} z_{vt}$. We now explicitly adopt a random coefficients framework to estimate the regression model,

$$y_{ivt} = \beta_{i0} + \beta_{iz} z_{vt} + \beta_{im} m_{iv} + \beta_{iw} z_{vt} m_{iv} + e_{ivt}. \tag{4.1}$$

In the potential outcomes framework of equations (3.1) and (3.2), $\beta_{iw} = \mu_1 - \mu_0 + \varepsilon_{1ivt} - \varepsilon_{0ivt}$. Under assumptions (A1), (A2), and (PPD), quantile treatment effects on the treated are identified by the coefficient, β_{iw}, on this interaction term at the corresponding quantiles of the outcome distribution.[11]

Results are presented in Figure 3. Given that we only observe 170 individuals after one year of treatment, the figure presents estimates for vigintiles (20-quantiles) of the outcome distribution among the treated. For ease of interpretation, we include a horizontal line at an impact of 230.8 kg, the sales volume required to repay the Abrabopa loan in the years studied. In the absence of any other costs of complementary inputs, individuals make a profit if and only if they experience a treatment effect in excess of this point.

Figure 3. Quantile treatment effects



Notes: Figure illustrates estimated quantile treatment effects and associated bootstrap (90%) confidence intervals. Non-parametric, block bootstrap confidence interval based on 400 repetitions, re-sampling at village level. Horizontal line at 230.8 kg indicates cocoa output required to repay direct cost of inputs.

Confidence intervals are estimated by a block bootstrap, with resampling conducted at the village-year level to account for potential non-independence in outcomes within these sampling units. Although we

[11] For quantile q, this is the treatment effect on treated individual i such that $i = \arg\inf_{y_{1i}} F_1(y_1 | w_{ivt} = 1) > q$.

can reject the null hypothesis of a treatment effect of zero for all but tails of the distribution, we fail to reject the hypothesis of a treatment effect above the break-even point for individuals at or below the 25th percentile and above the 80th percentile. (It should be noted that estimates at the upper end of the distribution are relatively imprecise, due to a small number of high-output farmers in our sample.)

Figure 3 suggests that, in this context, an exclusive emphasis on mean impacts misses important features of the distribution. In economic terms, there appears to be substantial heterogeneity in returns to program participation; in general, returns appear to be higher for more productive farmers. Crucially, we fail to reject the hypothesis of zero returns, net of repayment costs, for 7 of the 20 vigintiles of the treated population.

Both time-varying and time-invariant heterogeneity in returns may matter for the sustained use of a new technology. Even shocks that farmers know to be transient can affect future technology use, if they impact on farmers' willingness to take on risk (Dercon & Christiaensen 2007). On the other hand, time-invariant heterogeneity may affect farmers' willingness to continue through a process of experimentation – so long as farmers are uncertain about these returns prior to membership.[12] We address the questions of whether observed heterogeneity in returns can help to explain program retention, and through what mechanism, in the following section.

5 Treatment effect heterogeneity and program retention

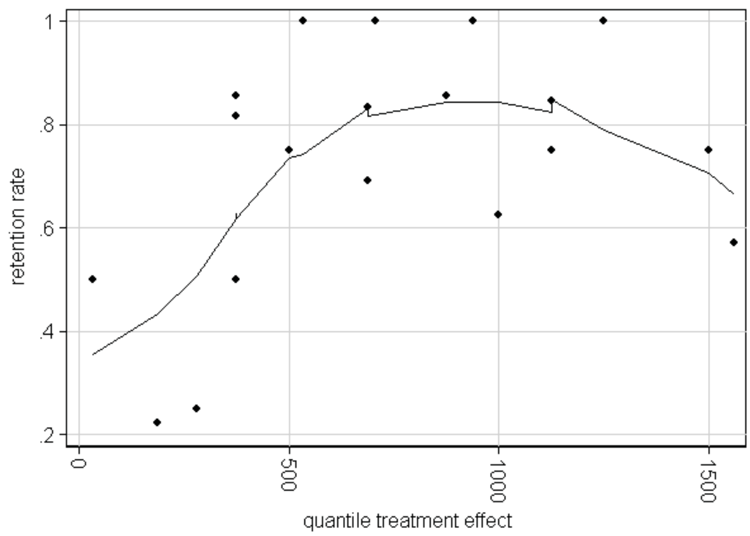
Treatment effects at each quantile of the outcome distribution are correlated with the program retention rate. We show this in three steps. First, we show that quantile treatment effects are correlated with program retention among individuals within a radius of that quantile. To take this further, we would ideally have a measure of individual-specific treatment effects, for the treated in our sample. We propose the two-period change in output among treated farmers as a proxy for the individual treatment effect. In our second step, we validate this measure of individual-specific impacts by showing that quantile treatment effects correlate with the two-period changes in output among individuals within a neighborhood of that quantile. Third, we employ this measure of individual-specific returns to show that this is correlated with retention rates, even after conditioning on a vector of own and group characteristics. This suggests that heterogeneity in ex post returns is relevant for the subsequent membership decisions of farmers, and that this relationship contains information beyond that in subgroups defined by observable characteristics.

To assess whether quantile treatment effects correlate with retention rates at that quantile, we must construct an estimate of the relevant retention rate. To do so, we take advantage of information from surrounding observations as follows: sort treated individuals by their outcomes. Suppose we are estimating treatment effects for Q quantiles: this will result in estimates for the $100/Q$ percentile, the

[12] In a variety of contexts, economic returns are highly sensitive to the precise quantities and combinations of inputs used. For example, Duflo and co-authors find that the right combination of fertilizer is highly profitable, but find that the Kenya Ministry of Agriculture's recommended dosage – in spite of having the highest agronomic return – actually results in farmers making a loss on average (Duflo et al. 2008)

$2 * 100/Q$ percentile, and so on. For quantiles $q = 1, .., Q$, we obtain an estimate of the retention rate at the corresponding $\frac{100q}{Q}$ percentile, by taking the mean retention rate between percentiles $\left(\frac{100q}{Q} - \frac{100}{2Q}, \frac{100q}{Q} + \frac{100}{2Q}\right)$.

Figure 4. Quantile treatment effects and program retention rates



Notes: Figure plots mean retention rates in the surrounding quantile against the treatment effect at that quantile of the outcome distribution. 20 quantiles are employed in estimation, leading to estimates of the QTE ate 19 interior points.

Figure 4 plots the relationship between quantile treatment effects and retention rates for 20-quantiles. As expected, there appears to be a positive relationship between the two. The rank correlation (Spearman's rho) between the QTE and the retention rate, across quantiles of the distribution, is 0.38. We employ a block bootstrap to obtain a sampling distribution for Spearman's rho that can be used to test the null hypothesis of independence between the two statistics. This allows us to account for potential non-independence of observations within villages. Given the relatively small number of clusters and the fact that the use of 20 quantiles throws away some of the variation in the data, we are unable to reject this null hypothesis: we obtain a p-value of 0.18 with 400 bootstrap repetitions.

To provide a more powerful test of the relationship between the magnitude of the treatment effect experienced and the retention decision, we construct a proxy for treatment effects at the individual level, and show that this individual measure correlates with retention, even after conditioning on possible confounding factors. Of course, since each farmer can make one and only one membership decision in a given period, the treatment effect that she experiences is never directly observable. This is the core of the "fundamental problem of causal inference" (Holland 1986).

Under certain (necessarily untestable) assumptions, individual measures of the actual treatment effect can be constructed. Our aim here is more limited: we seek to construct a proxy for the realized treatment effect, and to argue that variation in this proxy across individuals is correlated with variation in the true treatment effect. The purpose of this is to test the hypothesis that the decision to renew membership is affected by the realization of current output. We will test this using a binary choice model of the form

$$w_{iv,t+1}^{*} = \phi_0 + \phi_\tau \tau_{ivt} + \phi_x x_{ivt} + u_{iv,t+1}, \tag{5.1}$$

where membership in period $t+1$ is chosen if $w_{ivt}^{*} \geq 0$, τ_{ivt} is individual i's realized return in period t, and x_{ivt} represents a vector of controls for potential confounding factors, to be described below. Under the assumption of normality of the error term $u_{iv,t+1}$, the parameters of (5.1) can be estimated as a probit.

Because the individual's idiosyncratic return, τ_{ivt}, is not directly observable, we must estimate equation (5.1) with a proxy measure of the idiosyncratic return. For each individual in the treated group, we use two-period changes in cocoa output, $\hat{y}_{ivt} = y_{ivt} - y_{iv,t-2}$, as the basis for such a proxy for the experienced treatment effect.[13] Following the notation of equations (3.1) and (3.2), this will correspond to the true treatment effect, τ_{ivt}, only if $\lambda_{vt} = \lambda_{v,t-2}$ and $\varepsilon_{0ivt} = \varepsilon_{0iv,t-2}$. This would require both that any village-level components of production are constant over time, and that individuals' idiosyncratic output in the absence of Abrabopa is the same in the pre-exposure period $t-2$ as it is in the first period of exposure, t.

Neither of these assumptions will hold in practice: there will be village-level characteristics (such as rainfall) that vary across years, as well as shocks to potential output in the absence of fertilizer. Consequently the proxy $\hat{y}_{ivt}$ will be composed of the true treatment effect for individual i in period t, $\tau_{ivt} = y_{1ivt} - y_{0ivt}$, plus two terms that reflect the failure of these assumptions:

$$\hat{y}_{ivt} = \tau_{ivt} + \underbrace{(\lambda_{vt} - \lambda_{v,t-2})}_{\hat{\lambda}_{vt}} + \underbrace{(\varepsilon_{0ivt} - \varepsilon_{0iv,t-2})}_{\hat{\varepsilon}_{0ivt}}. \tag{5.2}$$

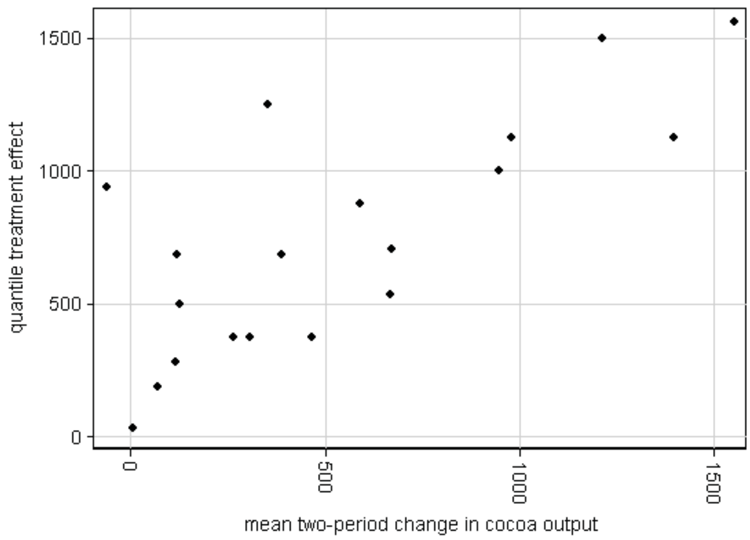
The correlation between this proxy and future membership decisions, $w_{iv,t+1}$, will reflect the effect of realized returns on individual decisions, as well as any correlation between the terms $\hat{\lambda}_{vt}$ and $\hat{\varepsilon}_{0ivt}$ and other determinants of membership. Thus measurement error in the proxy $\hat{y}_{ivt}$ creates a potential source of bias beyond the threat of omitted factors correlated with both the true, idiosyncratic treatment effect, τ_{ivt}, and the sustained adoption decision. We take up these considerations below.

Before turning to estimation of equation (5.1), we validate this measure by comparing the quantile treatment effects estimated in Figure 3 to the mean value of $\hat{y}_{ivt}$ for individuals in a neighborhood of the same vigintile, as was done with retention rates above. For example, we compare the quantile treatment effect at the 5th percentile of the outcome distribution among the treated, $F_1(y_{ivt})$, to the mean value of the individual measure $\hat{y}_{ivt}$ between the 2.5th and the 7.5th percentiles of the treated

[13] The advantage of using two-period changes, rather than one-period changes, as predictors is that if lagged output realizations affect current adoption decisions (Foster & Rosenzweig 2010), the latter would be a biased estimate of the ATT. However, we have replicated the results with the one-period change, and they are not substantially affected.

population. If the perfect positive dependence assumption holds – so that quantile treatment effects can be interpreted as the average response at that quantile – and if our proxy is a relatively precise estimate of the individual treatment effect, then the two measures should correlate closely. Figure 5 shows the relationship between these outcome measures for vigintiles of the outcome distribution among the treated. The measures of these two measures of impact are indeed closely related; they have a correlation coefficient of 0.76.

Figure 5. Validating alternative measures of impact: quantile treatment effects versus mean two-period changes



Notes: Horizontal axis is defined two-period change in cocoa output among treated farmers. The local average of this measure within a neighborhood of a given vigintile (20-quantile) is plotted against the corresponding quantile treatment effect.

Given the evidence to support this measure, $\hat{y}_{ivt}$, of individual treatment effects, we turn to estimation of equation (5.1). Probit marginal effects are presented in Table 4, which restricts attention to farmers who have joined Abrabopa in the first year of exposure in their village. The dependent variable in all specifications is the farmer's membership decision in the second year of exposure, which we observe through administrative records. In column (1), we demonstrate a positive relationship between a farmer's gain in output over the two-year period ending in their first year of membership and their membership decision in the second year. This effect is substantial. The standard deviation of the two-period change in output among the estimating sample of those who join Abrabopa is approximately 816 kg. This implies that a one standard deviation increase in this measure of the treatment effect is associated with a 13 percent increase in the likelihood of remaining in the program.

As mentioned above, this statistical association may be a biased estimate of the causal effect of (shocks to) individual returns to adoption on the membership renewal decision either because the proxy for individual returns employed introduces other factors correlated with the error term $u_{iv,t+1}$, or because the true idiosyncratic treatment effect is correlated with other (possibly unobserved) determinants of second-year membership decisions. We can address one aspect of this – namely, the bias that would arise if the change in village-level productivity, $\hat{\lambda}_{vt}$, is correlated with the unobserved component of individual choices, $u_{iv,t+1}$ – by including the mean two-period change in output among those who did not join Abrabopa in village v as a proxy for any such village-level shock. To do so we construct this mean as

$$\hat{y}^0_{ivt} = \frac{1}{N_{0vt}} \sum_{\{i:m_{iv}=0\}} y_{ivt} - y_{iv,t-2} \tag{5.3}$$

where N_{0vt} is the number of first-year non-adopters in village v and year t in our sample. As shown in column (2) of Table 4, the point estimate of the marginal effect of individual i's return on her subsequent membership decision is unaffected by inclusion of this variable. This suggests that the term $\hat{\lambda}_{vt}$ is unlikely to be a substantial source of bias in our proxy for individual returns.

Table 4. Individual output gains and program retention

	(1)	(2)	(3)	(4)	(5)
two-period change in output ($\hat{y}_{ivt}$)	0.161**	0.158***	0.146**	0.185**	0.354**
	(0.066)	(0.058)	(0.071)	(0.087)	(0.151)
village v non-adopters' mean change in output ($\hat{y}^0_{ivt}$)		0.045	0.072	0.068	0.132
		(0.214)	(0.189)	(0.189)	(0.215)
1{respondent failed to repay}			-0.703***	-0.720***	-0.810***
			(0.061)	(0.062)	(0.043)
1{other member failed to repay}			-0.449**	-0.415*	-0.312
			(0.204)	(0.216)	(0.283)
cocoa, kg (y_{ivt})				-0.048	-0.005
				(0.035)	(0.055)
Observations	140	140	107	107	81

Probit marginal effects reported, evaluated at mean values. Outcome is a dummy variable equal to 1 if first-year member in year t continued membership in year $t+1$. Output variables y_{ivt}, y^0_{ivt}, and y_{ivt} rescaled by dividing by 1,000 prior to estimation. All specifications include dummy variables for survey round. Column (5) contains controls for age, gender, education, household size, and farm size. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Given the group liability structure of the Cocoa Abrabopa Association, one might worry that the apparent effect of idiosyncratic returns on membership renewal is driven by repayment failures on the part of the member herself or by another member of her group. In such a case, it is Abrabopa's policy – albeit imperfectly enforced – to exclude all group members for one season. This is a confounding factor for our desired test of the hypothesis that farmers are learning from realizations of idiosyncratic returns, and making decisions accordingly.

To address this concern, in column (3) of Table 4, we include dummy variables that take a value of one if the member herself or any of her peers failed to repay. Unsurprisingly, both of these are highly significant. Failure to repay, conditional on the change in output, results in a further 70 percent decrease in the likelihood of renewed membership.[14] Repayment failures by one or more fellow members of the same group (excluding the individual under consideration herself) are associated with a further decline in the likelihood of sustained membership of approximately 45 percent. However, the inclusion of these additional controls does not substantially alter the point estimate of the association between our measures of treatment effects and retention rates.

There are at least two mechanisms through which realized returns could affect future adoption decisions. On the one hand, if returns vary across farmers and are not known with certainty, farmers might update their beliefs about their idiosyncratic returns to adoption on the basis of these realizations. This mechanism we refer to as *persistent* heterogeneity. On the other hand, as emphasized by Dercon and Christiaensen, poor realizations of returns might deter farmers from undertaking risky investments in the following periods, simply because they deplete farmers' buffer stocks and so their ability to undertake risk-increasing investments (Dercon & Christiaensen 2007). This is one mechanism underlying Foster and Rosenzweig's assertion that, when credit and insurance markets are imperfect, "lagged shocks to profits affect current input choices" (Foster & Rosenzweig 2010). Dercon and Christiaensen's mechanism links realized returns to subsequent adoption decisions even if all farmers returns are drawn from the same distribution. These two mechanisms have very different policy implications: whereas persistent heterogeneity implies that adoption is not optimal for some farmers, the buffer-stock mechanism suggests that interventions insuring farmers against some of the risks associated with technology adoption would encourage (welfare improving) adoption.

Notice that the buffer-stock mechanism does not relate the return to adoption, per se, with subsequent decisions. Instead it is the *level* of income in one period that effects the adoption decision. In a pure version of this story, where all farmers have the same return to adoption, the counterfactual outcome is irrelevant. Stronger yet, all past realizations of output – by virtue of helping the accumulation of buffer stocks – should be positively correlated with future adoption decisions. This suggests a simple test between the two mechanisms: to simply augment the specification in equation (5.1) with the current level of income. Conditional on current income, if the difference $\hat{y}_{ivt}$ remains positively associated with adoption decisions, this lends support to the hypothesis of learning about persistent heterogeneity. This is confirmed in column (4) of Table 4. While there is no significant relationship between current output and subsequent adoption, the association between the difference, $\hat{y}_{ivt}$, and adoption in the following year remains large and statistically significant.

Finally, in column (5) we address the concern that idiosyncratic returns to *hi tech* adoption, τ_{ivt}, may themselves be correlated with other variables that have a direct effect on the choice to sustain adoption. We do so by including controls for farm size, as well as the farmer's age, gender, education, and household size. Although data availability reduces the number of observations, we find no evidence that the correlation between idiosyncratic returns and these observed characteristics drives the evident association between returns and sustained adoption choices.

[14] The mean change in output levels among those who fail to repay is 408 kg less than among those who do repay; this difference is statistically significant at the 99 percent level.

From these results, we conclude that the data are consistent with a model in which individuals experience heterogeneous returns to participation in the Abrabopa program, and they condition their decisions to remain within the program on these experienced returns.[15] The association between measures of returns and future membership decisions does not appear to be driven by the 'supply' of the program – although repayment failures do cause expulsion in accordance with the program's bylaws. Nor is this association driven by current income alone: those who experience greater gains upon joining the program are more likely to remain, even at a given level of current production.

6 Conclusions

In many developing countries, the persistent adoption of agricultural technologies with high average returns is believed to be one of the principal policy challenges. Experimental and observational studies documenting these high average returns have led to a puzzle: why, if returns are so high, do farmers not adopt – and sustain – the use of agricultural technologies such as fertilizer or hybrid seeds?

We have presented evidence consistent with the view that high average returns mask substantial, persistent heterogeneity in realized returns. Farmers at the low end of the distribution of cocoa production exhibit particularly low returns – so much so that we are unable to reject a zero economic return for the bottom quarter of the distribution. We have shown evidence to support the view that this heterogeneity matters, economically speaking: farmers exhibiting low returns, by various measures, are less likely to continue use of the technology.

This ex post heterogeneity likely reflects a combination of persistent heterogeneity and time-varying riskiness of returns. Both may have consequences for decisions to sustain or to disadopt a given technology. Farmers may experiment in order to learn about their specific, time-varying component. Time-varying shocks to these returns may lead Bayesian farmers astray. But they may also affect adoption decisions directly, either because they lead to non-repayment of loans, or because they force farmers to deplete buffer stocks of savings, or social or other forms of collateral, in order to repay. In our data and context, there is some evidence that the latter effect is not driving the relationship between realized returns and disadoption: our measure of realized returns retains its statistical and economic significance even after conditioning on the level of past output.

The distinction between the effect of transient and persistent heterogeneity in returns seems a valuable area for future work. If persistent heterogeneity is quantitatively important, policymakers will need to be cautious in advocating widespread adoption of such technologies. Even when average returns are high, many farmers may stand to lose.

[15] Sustained membership in Abrabopa is not equivalent to sustained use of the technology it provides. Examination of whether treatment heterogeneity correlates with continued use of the *hi-tech* package is not possible in this context, since we observe only whether broad categories of inputs such as fertilizer were used, but not their exact make or proportions on a given plot; moreover, we only have these production data for 29 of the first-year members in our sample who subsequently drop out of Abrabopa. Among these 29 farmers, 14 report using fertilizer in some form and quantity.

References

Abbring, J. H. & Heckman, J. J. (2007), Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation, *in* J. J. Heckman & E. E. Leamer, eds, 'Handbook of Econometrics', Vol. 6, Elsevier, chapter 72, pp. 5143–5303.

Bandiera, O. & Rasul, I. (2006), 'Social networks and technology adoption in northern mozambique', *Economic Journal* **116**(514), 869–902.

Bitler, M. P., Gelbach, J. B. & Hoynes, H. W. (2006), 'What mean impacts miss: Distributional effects of welfare reform experiments', *American Economic Review* **96**(4), 988–1012.

Cocoa Abrabopa Association (2009), 'Developments and challenges of the CAA programme', Presentation to workshop on "Improving Productivity among Ghanaian Cocoa Farmers through Group Lending", Accra, Ghana.

Conley, T. G. & Udry, C. R. (2010), 'Learning about a new technology: Pineapple in Ghana', *AER* **100**, 35–69.

Dammert, A. C. (2009), 'Heterogeneous impacts of conditional cash transfers: Evidence from nicaragua', *Economic Development and Cultural Change* **58**, 53–83.

Dercon, S. & Christiaensen, L. (2007), 'Consumption risk, technology adoption and poverty traps: evidence from Ethiopia', World Bank Policy Research Workign Paper 4257.

Djebbari, H. & Smith, J. (2008), 'Heterogeneous impacts in PROGRESA', *Journal of Econometrics* **145**(1–2), 64–80.

Duflo, E. (2003), 'Poor but rational?', Working paper, MIT.

Duflo, E., Kremer, M. & Robinson, J. (2008), 'How high are rates of return to fertilizer? Evidence from field experiments in Kenya', *American Economic Review* **98**(2), 482–488.

Duflo, E., Kremer, M. & Robinson, J. (2009), 'Nudging farmers to use fertilizer: Theory and experimental evidence from kenya', NBER Working Paper No. 15131.

Field, E. (2005), 'Property rights and investment in urban slums', *Journal of the European Economic Association* **3**(2/3), 279–290.

Foster, A. D. & Rosenzweig, M. R. (1995), 'Learning by doing and learning from others: Human capital and technical change in agriculture', *The Journal of Political Economy* **103**(6), 1176–1209.

Foster, A. D. & Rosenzweig, M. R. (2010), 'Microeconomics of technology adoption', Yale University, Economic Growth Center, Center Discussion Paper No. 984.

Frêchet, M. (1951), ‘Sur les tableaux de corrélation dont les marges sont données’, *Annals University Lyon: Series A* **14**, 53–77.

Heckman, J. J. (2010), ‘Building bridges between structural and program evaluation approaches to estimating policy’, *Journal of Economic Literature* **48**(2), 356–398.

Heckman, J. J. & Hotz, V. J. (1989), ‘Alternative methods for evaluating the impact of training programs’, *Journal of the American Statistical Association* **84**, 862–874.

Heckman, J. J. & Navarro-Lozano, S. (2004), ‘Using matching, instrumental variables, and control functions to estimate economic choice models’, *Review of Economics and Statistics* **86**, 30–57.

Heckman, J. J., Smith, J. & Clements, N. (1997), ‘Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in program impacts’, *Review of Economic Studies* **64**(4), 487–535.

Hoeffding, W. (1940), ‘Maastabinvariate korrelationstheorie’, *Schriften des Mathematischen Instituts und des Institutes Fur Angewandte Mathematik der Universitat Berlin* **5**, 179–223.

Holland, P. W. (1986), ‘Statistics and causal inference’, *Journal of the American Statistical Association* **81**(396), 945–960.

Jovanovic, B. & Nyarko, Y. (1996), ‘Learning by doing and the choice of technology’, *Econometrica* **64**(6), 1299–1310.

Moser, C. M. & Barrett, C. B. (2006), ‘The complex dynamics of smallholder technology adoption: the case of SRI in madagascar’, *Agricultural Economics* **35**, 373–388.

Munshi, K. (2004), ‘Social learning in a heterogeneous population: technology diffusion in the indian green revolution’, *Journal of Development Economics* **73**(1), 185–213.

Opoku, E., Dzene, R., Caria, S., Teal, F. & Zeitlin, A. (2009), ‘Improving productivity through group lending: Report on the impact evaluation of the Cocoa Abrabopa Initiative’, Centre for the Study of African Economies, technical report no. REP2008-01.

Pearl, J. (2009), ‘Causal inference in statistics: an overview’, *Statistics Surveys* **3**, 96–146.

Suri, T. (2007), ‘Selection and comparative advantage in technology adoption’, Unpublished, MIT.

World Bank (2008), *World Development Report 2008: Agriculture for Development*, The International Bank for Reconstruction and Development, Washington, D.C.

Zerfu, D. & Larson, D. F. (2010), ‘Incomplete markets and fertilizer use’, World Bank, Policy Resaerch Working Paper No. 5235.

Appendix A. Tables

Table A1. Characteristics of sampled farmers

	Survey wave: 2008				Survey wave: 2009			
	Villages reached in 2007		Villages reached in 2008		Villages reached in 2008		Villages reached in 2009	
	Member	Non-member	Member	Non-member	Member	Non-member	Member	Non-member
cocoa, kg.	1840.94	1395.87	1615.20	1460.00	1606.47	1144.40	1034.18	1376.08
	(1867.80)	(1976.24)	(2098.11)	(1344.95)	(1362.87)	(1072.45)	(1152.82)	(1142.40)
land devoted to cocoa, ha.	4.35	6.43	3.29	4.16	3.48	4.18	3.39	3.99
	(3.73)	(10.94)	(2.40)	(4.03)	(2.59)	(2.97)	(3.32)	(3.54)
age	51.82	50.12	49.99	47.23	50.37	49.00	44.58	45.54
	(12.46)	(11.71)	(13.65)	(11.14)	(13.33)	(10.11)	(13.05)	(16.05)
female	0.14	0.24	0.24	0.30	0.25	0.27	0.20	0.15
	(0.35)	(0.44)	(0.43)	(0.47)	(0.44)	(0.45)	(0.40)	(0.37)
adults in household	3.22	2.52	3.04	3.33	3.09	3.08	3.48	3.77
	(1.72)	(1.50)	(1.98)	(2.48)	(2.01)	(1.44)	(2.05)	(2.12)
farmer education: Junior Secondary School or above	0.68	0.60	0.59	0.47	0.58	0.50	0.56	0.54
	(0.47)	(0.50)	(0.50)	(0.51)	(0.50)	(0.51)	(0.50)	(0.51)
any fertilizer used	0.99	0.44	0.31	0.19	0.97	0.48	0.46	0.46
	0.11	0.5	0.47	0.4	0.17	0.51	0.5	0.51
fertilizer, 50 kg. Bags	8.49	4.23	2.84	2.00	11.46	7.72	4.38	7.38
	(6.88)	(8.24)	(5.25)	(5.44)	(13.77)	(12.05)	(6.26)	(10.97)
household adult labor: days worked on farm	87.35	88.97	56.19	63.17	62.39	75.59	72.87	99.68
	(86.67)	(82.25)	(74.00)	(61.68)	(70.14)	(74.47)	(72.14)	(147.72)
paid labor: days worked on farm	36.24	31.03	41.56	37.36	37.42	24.41	17.48	26.95
	(55.98)	(52.09)	(63.68)	(55.88)	(61.16)	(49.87)	(31.12)	(41.32)
nnoboa labor: days worked on farm	3.77	3.19	2.14	1.17	3.04	0.34	3.83	2.03
	(7.27)	(7.99)	(5.97)	(2.58)	(7.26)	(0.86)	(10.02)	(6.28)
number of observations	82	41	88	42	72	29	95	37
number of villages	8		8		8		8	

Notes: (1) Member/non-member refers to membership decision in year village first reached by CAA. (2) Number of observations is the number of observations with complete output data, by survey round and membership status. (3) Villages reached in 2008 by CAA were surveyed in both 2008 and 2009. (3) (Pseudo) treatment effect is estimated using the pooled cross section, difference-in-differences specification of Table 2, column (2); p-values shown in parentheses. (4) A Wald test that pseudo-treatment effects are equal to zero for all predetermined characteristics (farm size, age, gender, adults, and education) yields a chi-squared statistic of 3.86, with an associated p-value of 0.57 (5 degrees of freedom).