DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Learning Through Noticing: Theory and Experimental Evidence in Farming

Citation

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2012. Learning Through Noticing: Theory and Experimental Evidence in Farming. HKS Faculty Research Working Paper Series and CID Working Papers (RWP12-044 and 245), HKS Faculty Research Working Paper Series.

Published Version

http://web.hks.harvard.edu/publications/workingpapers/citation.aspx?PubId=8586

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:9804491

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

HARVARD Kennedy School
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

Learning Through Noticing: Theory and Experimental Evidence in Farming

Faculty Research Working Paper Series

Rema Hanna

Harvard Kennedy School, NBER and BREAD

Sendhil Mullainathan

Harvard University and BREAD

Joshua Schwartzstein

Dartmouth College

October 2012
RWP12-044

Visit the **HKS Faculty Research Working Paper** series at:
http://web.hks.harvard.edu/publications

The views expressed in the **HKS Faculty Research Working Paper Series** are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.

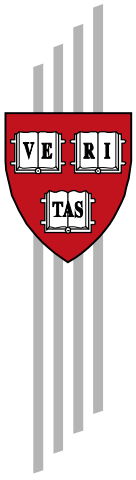
www.hks.harvard.edu

Learning Through Noticing: Theory and Experimental Evidence in Farming

Rema Hanna/Sendhil Mullainathan/Josh Schwartstein

CID Working Paper No. 245
September 2012

© Copyright 2012 Rema Hanna, Sendhil Mullainathan, Josh Schwartstein, and the President and Fellows of Harvard College



Working Papers

Center for International Development at Harvard University

Learning Through Noticing: Theory and Experimental Evidence in Farming*

Rema Hanna
Harvard University, NBER and BREAD

Sendhil Mullainathan
Harvard University and BREAD

Joshua Schwartzstein
Dartmouth College

September 6, 2012

Abstract

Existing learning models attribute failures to learn to a lack of data. We model a different barrier. Given the large number of dimensions one could focus on when using a technology, people may fail to learn because they failed to notice important features of the data they possess. We conduct a field experiment with seaweed farmers to test a model of "learning through noticing". We find evidence of a failure to notice: On some dimensions, farmers do not even know the value of their own input. Interestingly, trials show that these dimensions are the ones that farmers fail to optimize. Furthermore, consistent with the model, we find that simply having access to the experimental data does not induce learning. Instead, farmers change behavior only when presented with summaries that highlight the overlooked dimensions. We also draw out the implications of learning through noticing for technology adoption, agricultural extension, and the meaning of human capital.

*We thank Andy Newman, Matthew Rabin, and Andrei Shleifer for helpful comments, as well as seminar participants at the Behavioural Decision Theory Conference, Berkeley, BU, and Harvard/MIT.

1 Introduction

Learning models suggest several reasons why people may fail to operate on the productivity frontier. Experimentation can impose both real costs and opportunity costs. Even with experimentation, updating may be slow because outcomes are noisy (Besley and Case 1993, 1994). Data from neighbors' experiences is only partly helpful: their contexts may differ, and some technologies may require actual use in order to learn (Jovanovic and Nyarko 1996; Foster and Rosenzweig 1995; Conley and Udry 2010). In all these cases, the bottleneck is data: people fail to learn because because they lack sufficient data to learn from. In this paper, we highlight another reason why people fail to learn.

Imagine having dinner at a friend's house, where you watch him cook steaks using a cast iron skillet. These steaks are so good that you buy a skillet and decide to try making them yourself. Your friend's kitchen and cooking skills are simliar enough to yours that, under traditional learning models, you should end up with similar steaks. Sure enough, the first batch of steaks is delicious. The second, though, is not as good and they get worse after that. What happened? You failed to notice something important. When washing his skillet, your friend did not use any soap. Washing the skillet with soap removes the "seasoning" and results in a sticky surface that cooks unevenly and produces less flavor. Not noticing this, you washed your skillet before cooking. This failure to notice was not by chance. Watching your friend cook, there was an endless array of details you could have paid attention to, from the temparature in the room to the color of his apron. You decided which details to encode or "notice" based on your beliefs. Since you did not believe washing (or lack thereof) had anything to do with cooking, you did not pay attention to that.[1] There was no lack of data, but a failure to notice something in that data. We call this *learning through noticing*.

Some practical examples also match this pattern of learning through noticing. For many years, doctors had the data they needed to prevent operating room infections. But the importance of a

[1] If your friend watched you attempt the recipe, she could point out your error. However, it is not the case that explicit communication always bridges this gap. For example, she would need to know your tendency not to notice certain details and to tailor her instructions accordingly.

sterile operating room was not recognized until the germ theory of disease spread (Nuland 2003, Gawande 2004). Until then, doctors did not learn that they themselves were the source of infections. Similarly, recent evidence suggests that Indian textile manufacturers fail to adopt key management practices, such as keeping the factory floor free of waste, despite exposure to natural variation that points to the importance of those practices (Bloom et al. 2012). Other work suggests that shopkeepers in Kenya fail to attend to the fact that keeping a hefty stock of small change can improve their profitability (Beaman et al. 2012).

We use Schwartzsteins's (2012) model of selective attention to build a theory of learning through noticing in the context of technology adoption and use. Agents in the model wrestle with numerous dimensions in learning how to use a technology. A farmer, for example, must decide on the timing of planting, the depth and spacing of seeds, the amount of fertilizer to use, when to use it and so on.[2] In addition to these decisions, the agent in our model must also decide which dimensions to attend to, where attending is costly. The agent is "Savage rational": he updates and chooses in a Bayesian way given prior beliefs and the costs of paying attention. The model highlights a feedback loop. An agent who believes that a particular dimension does not matter may not focus on it. As such, he may never process freely available data that could contradict his false belief. This generates an empirical prediction: a failure to learn comes from a failure to notice. So learning failures should be concentrated on dimensions where agents report ignorance, i.e., where they cannot answer key questions about what they themselves do (or have done) along that dimension.

We test the model's predictions using data from a field experiment with seaweed farmers. Seaweed is farmed by attaching strands ("pods") to lines submerged in the ocean. As in the model, a large number of dimensions might affect yield.[3] We focus on the spacing between the pods and

[2]While the model presented in this paper is specific to issues related to technology adoption and use, Schwartzstein (2012) presents a general model of belief formation when an agent is selectively attentive. For other more general approaches to modeling limited attention in economic settings, see Sims (2003), Bordalo et al. (2012), Koszegi and Szeidl (2011) and Gabaix (2011).

[3]For example, the strength of the tide, the time of day, the temperature, the tightness with which pods are attached, the strain of pods used and many other dimensions could matter. This is a persistent issue to keep in mind. We focus on two dimensions for parsimony. Actual demands on attention are larger than this narrative suggests since these are two of many dimensions that all must be attended to (or not).

the size of the pods. To look for failures to notice, we directly asked farmers what they know about their own production techniques and production functions. As one might expect, farmers are quite knowledgeable about many dimensions. However, they do report ignorance on one key dimension: they do not know the size of pods they use (and few will even venture a guess about what the optimal size should be). To test whether this failure to notice translates into a failure to learn, we conducted experiments on farmers' own plots, varying both pod size and pod spacing. On pod spacing – which farmers tended to notice – our findings suggest they were close to the optimum. In contrast, on pod size – which farmers did not notice – our findings suggest they were far from the optimum.

To more sharply test the model, we examine farmers' response to the trial. The model predicts that, even though the trials generate data to learn from, simply participating in them should not affect pod size. The intuition is that the farmers' own behavior generated an experiment of sorts every season – their failure to notice size created random variation in pod sizing. In effect, the trial presents farmers with similar data that they already had access to, but did not notice. Consistent with this prediction, we find little change in farmers' behavior following participation in the trial. The model further suggests a way to induce learning: providing a summary of the data that explicitly highlights neglected dimensions should affect behavior. Consistent with this prediction, we find that farmers changed their production methods after we presented them with the trial data on yield broken down by pod size (from their own plot).

The analysis described so far focuses on *how* to use a technology. We extend the model to study the decision of *whether* to use a technology. A key result here is that agents can experience predictable disappointment. Agents can observe data – perhaps in a demonstration – that leads them to believe that they can use a technology profitably. Yet, when they adopt it, they systematically earn negative surplus in a way that is predictable to an outsider who has knowledge of the production function and what agents attend to. Disappointment happens –as in the skillet example– when agents do not notice all of the dimensions necessary for effective use. We re-interpret some existing evidence on technology adoption in this light, prominently the fertilizer experiments by Duflo,

Kremer and Robinson (2008). This logic also provides a different perspective on agricultural extension activities where farmers are shown how to use technologies. In our model, the effectiveness of extension depends on whether the agents' priors lead them to notice the relevant features of a technology. We call a technology "filter congruent" if agents' beliefs on what to notice match up well with the dimensions needed to effectively use the technology. Non-filter-congruent technologies are prone to disappointment. We generalize this concept to develop a broader understanding of human capital. From a learning point of view, human capital is more than the accumulated knowledge of what has worked. It is also the accumulated understanding of *what to notice*.

The paper proceeds as follows. Section 2 presents the model of learning through noticing given a choice of technology and develops our experimental hypotheses. Section 3 describes the experimental design and data. Section 4 presents the results. Section 5 extends the model to consider the technology adoption decision and further applications. Section 6 concludes with a discussion of broader issues, for example how learning through noticing changes our understanding of the role of human capital, experience, and education in promoting the efficient use of technologies.

2 Basic Model

2.1 Setup

Consider a farmer whose learning problem is how to set inputs to maximize profits. While, in reality, there are many dimensions, we assume that the farmer knows his production technology up to a single input, $s \in \{1, 2, 3, \ldots, S\}$, in order to simplify the model. Over the course of two periods, $t = 1, 2$, the farmer collects data that may inform his choice of s. At a given period of time, the farmer chooses which level of s to use for each of N pieces of land. He can either measure to guarantee a particular level, which costs $m > 0$, or choose not to measure. If the farmer does not measure, then the input level is random and uniformly drawn across possible levels $\mathcal{S} = \{1, 2, \ldots, S\}$. To illustrate, in our seaweed application, the farmer may not know the optimal pod size. In every period, he has the opportunity to either choose particular sizes for each

of his N pods or not to precisely measure them, which would thus lead to random variation in pod sizes.[4]

Given input s, the yield from a piece of land p at time t is given by:

$$y_{pt} = f(s_{pt}) + \varepsilon_{pt}, \tag{1}$$

where $\varepsilon_{pt} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is a random shock that is independent across time and pieces of land. Denote the true production process (the true value of $f(\cdot)$) by $f_0(\cdot)$ and the value that maximizes this function by s_0 (for simplicity, if there are multiple such values, pick one). The farmer is initially uncertain whether and how the input affects the yield – that is, about $f_0(\cdot)$ – and updates his beliefs over time.

The farmer attaches prior weight $\pi \in (0, 1)$ to a model, P, in which s affects yield and weight $1 - \pi$ to a model, NP, in which s does not affect yield.[5] Under model P,

$$f(s) = \theta(s), \tag{2}$$

where each $\theta(s)$, $s \in \mathcal{S}$, is independently drawn from a $\mathcal{N}(\mu, \sigma^2)$ distribution. Under this S parameter model, yield almost surely varies in the input. On the other hand, under model NP,

$$f(s) = \theta, \tag{3}$$

where θ is drawn from a $\mathcal{N}(\mu, \sigma^2)$ distribution. Under this single parameter model, yield does not vary in the input. Combining Equations (1)-(3), under model P:

$$y_{pt} = \theta(s_{pt}) + \varepsilon_{pt} \tag{4}$$

[4]When the farmer does not measure, he also has the option of paying attention to the relationship between the input level and the yield. This choice variable will be discussed in detail after specifying the environment.

[5]To remember these labels, P is short-hand for "pod size matters" in the main application.

while under model NP,

$$y_{pt} = \theta + \varepsilon_{pt}. \tag{5}$$

Fully Attentive Benchmark

A fully attentive farmer updates his beliefs about whether and how the input affects yield using Bayes' rule, as applied to the first period history:

$$h = (y_{p1}, s_{p1})_{p=1}^{N}. \tag{6}$$

To illustrate, the standard updating formula for a normal learning model (e.g., DeGroot, 1970) implies that, given h, a farmer's posterior expectation of yield for a given level of s' is:

$$\begin{aligned} E[f(s')|h] &= \pi^{+} E[f(s')|h, P] + (1 - \pi^{+}) E[f(s')|h, NP] \\ &= \pi^{+} E[\theta(s')|h, P] + (1 - \pi^{+}) E[\theta|h, NP] \\ &= \pi^{+} \left(\frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + n(s')\sigma^2} \mu + \frac{n(s')\sigma^2}{\sigma_{\varepsilon}^2 + n(s')\sigma^2} \bar{y}_{s'} \right) + (1 - \pi^{+}) \left(\frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + N\sigma^2} \mu + \frac{N\sigma^2}{\sigma_{\varepsilon}^2 + N\sigma^2} \bar{y} \right), \end{aligned} \tag{7}$$

where π^{+} is the posterior probability placed on model P, $n(s')$ stands for the number of times the farmer used input level s' in the first period, $\bar{y}_{s'}$ is the average first-period yield across pieces of land with input level s', and $\bar{y}$ is the average first-period yield across all pieces of land.

Selective Attention

We follow Schwartzstein (2012) and make the additional assumption that the farmer must expend effort in order to attend to the input level relative to yield, so he is only selectively attentive. The farmer faces a cost $e \in (0, m)$ of keeping track of the input level applied to a given piece of land if it was not measured, and will do so only if the expected benefits from doing so, i.e., from being able to use this information when choosing the input level in the next period, exceed this cost. In

the second period, he updates his beliefs not based on the full history h, but on what he noticed, i.e., based on:

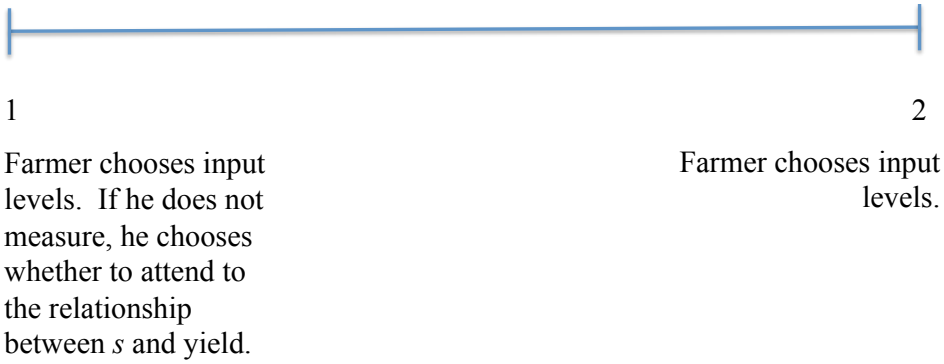
$$\hat{h} = (y_{p1}, \hat{s}_{p1})_{p=1}^{N}, \tag{8}$$

where

$$\hat{s}_{p1} = \begin{cases} s_{p1} & \text{if the farmer kept track of the input level applied to piece of land } p \\ \varnothing & \text{if the farmer did not keep track of the input level applied to piece of land } p. \end{cases} \tag{9}$$

2.1.1 Learning by Doing

The farmer is risk neutral and maximizes the expected undiscounted sum of yield across periods, net of the costs of measuring and attending to the input.[6] Timeline 1 summarizes the flow of activities.



Timeline 1: Learning by Doing

Proposition 1. *1. If: (i) π is sufficiently low or (ii) e and m are sufficiently large, then the farmer does not measure or attend to the input level applied to any piece of land in the first period. In this case, the farmer's posterior expectation of $f(s)$ does not vary in s, and he chooses not to measure in the second period.*

[6]Implicitly, we are making the simplifying assumption that the differential cost to the farmer of using different inputs is negligible.

2. *If: (i) π is sufficiently large (fixing sufficiently low e and m) or (ii) e and m are sufficiently low (fixing π), then the farmer attends to the input level applied to at least one piece of land in the first period. In this case, the farmer's posterior expectation of $f(s)$ almost surely varies in s so long as $\sigma_\varepsilon^2 > 0$. He either measures the input level with the highest associated posterior expectation of $f(s)$ in the second period or does not measure.*

Proof. See the Appendix A for all proofs. ■

Proposition 1 says that if the farmer does not place much initial weight on an input level, or the cost of measuring and attending to the input is sufficiently large, then he will not measure or attend to that input. Such a farmer will continue to be indifferent between input levels despite having access to readily available information about the relationship between the input level and yield. He will not know what he does nor will he change his practice with respect to that input over time. This prediction is independent of the underlying production function $f_0(\cdot)$: the farmer can make arbitarily large mistakes by failing to attend to the input, as an incorrect belief that an input is unimportant is self-confirming given selective attention (Schwartzstein 2012).[7] On the other hand, if the farmer places enough initial weight on the input and the cost of measuring and tracking that input is sufficiently low, then he will attend to the input in the first period. In this case, he will come to expect some levels to achieve greater yields than others and may measure the input in the second period.[8]

To observe the potential loss from the farmer not paying attention, consider the limiting situation where $\sigma_\varepsilon^2 = 0$, so that a single observation using input level s' is enough to identify $f_0(s')$, and where N is very large relative to S, so that a farmer is likely to learn the optimal input level s_0 when he does not measure but attends to the input level in the first period. In this case, when the farmer does not measure but attends in the first period he is likely to make the same choice in the second as if he were endowed with knowledge of the true production function (i.e., he will make

[7] This prediction is not shared by most economic models of inattention (e.g., Sims 2003 or Gabaix 2011), since they abstract from learning by assuming that agents are endowed with perfect knowledge of the mapping between inputs and payoff relevant outputs. In such models, agents do not attend to an input only if it is *truly* not very important.

[8] Statements regarding almost sure convergence are with respect to the true underlying distribution, i.e., conditional on $f_0(\cdot)$.

the "optimal" choice, s_0). This is not true when he does not attend in the first period. Supposing that $f_0(s_0) - m$ is large relative to $1/S \sum_{s'} f_0(s')$, not paying attention results in a substantial income loss. This generates the following testable predictions:

Prediction 1. Farmers may not attend to all dimensions of the production function.

Prediction 2. When farmers do not attend to an input dimension, choices along that dimension may be far from optimal. When farmers attend, choices are more likely to be optimal.

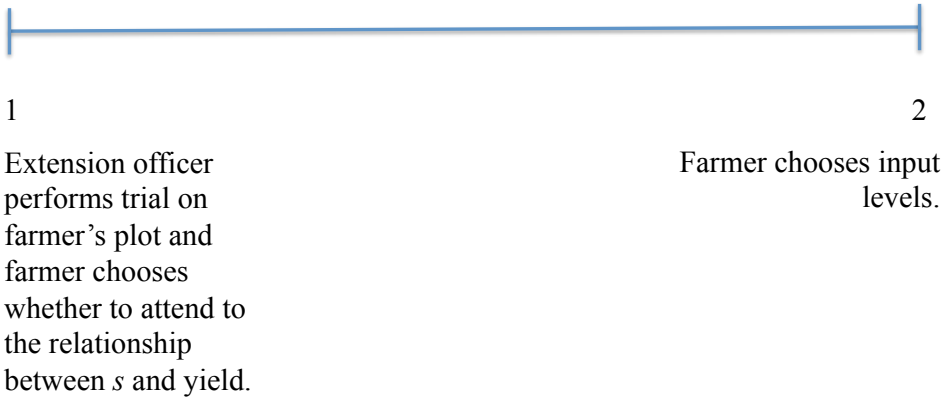
In addition to these predictions, the baseline model also illuminates methods to diagnose failures to notice and the resulting failures to learn. In particular, it suggests a simple way of testing whether farmers are attending to an input: elicit beliefs about current practices and best practices. When farmers *do not attend*, the model predicts that they will not precisely know their previously used input levels, nor will they have an opinion about the relationship between those levels and yield.

Finally, note that the model also implies that farmers are less likely to notice an input when the cost of attending to that input is higher or when the option value of paying attention is lower. Put differently, farmers are less likely to attend to an input when it is "more difficult" to learn its relationship to yield, e.g., when the relationship is noisy. Thus, one novel prediction of our model, relative to more standard learning models, is that such barriers may not just slow down the speed of learning, but shut it down entirely.

2.1.2 Learning from Others

In addition to learning through their own experience, the farmer could learn from observing others. Consider a useful benchmark: the farmer can observe someone who actively experiments with different input choices, sidestepping the issues that arise from insufficient experimentation. To match the main application, suppose the farmer learns from an experimental trial conducted by someone else, who we will label an extension officer, on the farmer's own plot. We continue to assume a two period model. In the first period, assuming that $N/S \equiv n$ is integer-valued, the

officer randomly assigns n pieces of land to input level $s = 1$, n to level $s = 2$, and so forth. The farmer's only choice in that period is whether to attend to the input level relative to yield. In the second period, the farmer chooses the input level. There are two regimes: one in which the officer presents the farmer with results from the experimental trial prior to his choice of input ("learning by summary") and one in which he does not ("learning by observing"). The farmer knows the regime he is in. We start by discussing the learning by observing regime which is summarized in Timeline 2.



Timeline 2: Learning by Observing

Proposition 2. *If the farmer would not have attended to any input level when learning by doing, then he also may not in the learning by observing regime: If π is sufficiently low or e and m are sufficiently large, then the farmer does not attend to the input level applied to any piece of land in the learning by observing regime. Such a farmer's posterior expectation of $f(s)$ does not vary in s, and he chooses not to measure the input in the second period.*

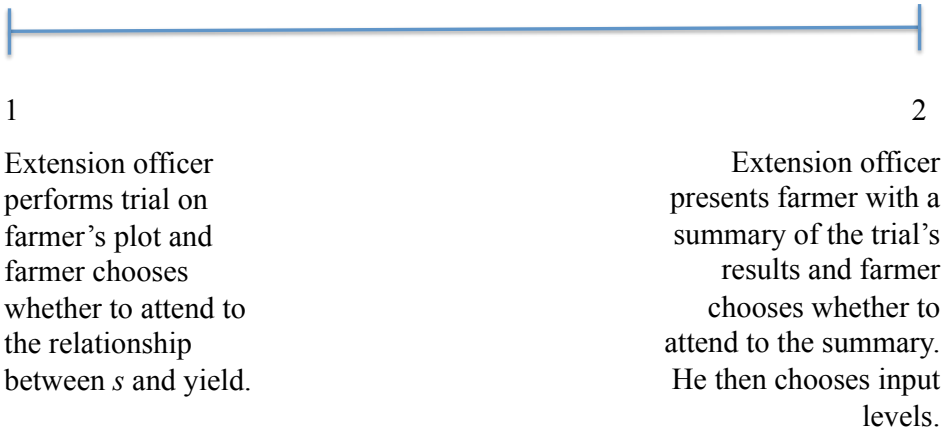
Proposition 2 says that when the farmer's failure to learn the relationship between the input level and yield is a consequence of selective attention, a demonstration that targets experimentation alone may be ineffective.[9] An inattentive farmer *already experiments*, albeit unintentionally, and does not learn the relationship between the input level and yield despite having access to freely

[9]This analysis implicitly assumes that the farmer does not update his prior as a result of his trial participation. There are reasons to believe this assumption held in our experimental setting, as discussed below.

available information which would aid in forming a belief about this relationship. Thus, exogenously providing such a farmer with this same information, in essentially the same way, may not change his beliefs.

Prediction 3. Farmers may not learn much from observing experimental trials that exogeneously vary inputs.

Adding Summary Policies that make it easier for the farmer to learn the empirical relationship between the input level and yield may be more effective. Consider a regime (the "learning by summary" regime that is described in Timeline 3) where, in addition to performing an experimental trial on a farmer's plot, the officer calculates and presents the farmer with detailed results from the trial prior to his second period input choice. Namely, the farmer is told the level s^* that achieved the greatest sample average yield, as well as the corresponding sample average $\bar{y}_{s^*}$. Since the cost of keeping track of the input relative to yield for a given piece of land is e, it is natural to assume that the cost of attending to $(\bar{y}_{s^*}, s^*)$ is also e.



Timeline 3: Learning by Summary

Under certain assumptions (e.g., $\sigma_\varepsilon^2 = 0$), $(\bar{y}, \bar{y}_{s^*}, s^*)$ is a sufficient statistic relative to the farmer's second-period decision (i.e., he will make the same choice whether he knows all of h or just $(\bar{y}, \bar{y}_{s^*}, s^*)$). This statement is formalized and proved in Appendix A. Under these assumptions, while the expected benefit of attending to the officer's recommendation is the same as that of

keeping track of the input applied to *each* piece of land, the cost is lower by a factor of N (the number of pieces of land). This observation lies at the root of the following Proposition.

Proposition 3. *A farmer is more likely to attend to information that helps him estimate the relationship between the input level and yield when he is presented with summary statistics of results from the experiment: Supposing $\sigma_\varepsilon^2 = 0$, a farmer who would not keep track of the input level applied to any piece of land in the learning by observing regime may attend to $(s^*, \bar{y}_{s^*})$ in the learning by summary regime; a farmer who would not attend to $(s^*, \bar{y}_{s^*})$ in the learning by summary regime would also not keep track of the input level applied to any piece of land in the learning by observing regime. A farmer who attends to $(s^*, \bar{y}_{s^*})$ either uses s^* in the second period or does not measure.*

Proposition 3 shows that highlighting features of the data that the farmer had seen but failed to notice can be quite useful. The fact that it is enough to present the farmer with $(\bar{y}_{s^*}, s^*)$ for these purposes (i.e., that $(\bar{y}, \bar{y}_{s^*}, s^*)$ is a sufficient statistic) is rather specific to assumptions about the farmer's prior as well as the environment. More generally, a selectively attentive farmer is more likely to learn from some sufficient statistic relative to his second period-decision so long as it is less costly to process this statistic than to attend to the input level applied to every piece of land.

Prediction 4. Farmers are more likely to learn from summaries of experimental trial results.

2.2 Summary of Predictions

The model suggests ways to diagnose and predict failures to notice, and yields the following predictions, which we test in our main application:

1. Farmers may not attend to all dimensions of the production function.

2. When farmers do not attend to an input dimension, choices along that dimension may be far from optimal. When farmers attend, their choices are more likely to be optimal.

3. Farmers may not learn much from observing experimental trials that exogeneously vary inputs.

4. Farmers are more likely to learn from summaries of trial results.

3 Experimental Design and Data

3.1 Setting

Our experiment takes place with seaweed farmers in the Nusa Penida district in the province of Bali, Indonesia. The farmers in this area have been growing seaweed since it was introduced in the early 1980s. Most farmers initally grew a variety called spinosim, but some have moved to a different strain called cottoni, due to some combination of buyer advice and government and NGO extension programs. While there are different production methods, most farmers in this area use the bottom method: in each plot, the farmers drive wooden stakes in the shallow bottom of the ocean, and then attach lines across the stakes. They then take raw seaweed from the last harvest and cut them into pods. The farmers then plant these pods at a given interval on the lines. The optimal pod size can be determined by numerous factors: for example, bigger seedlings may result in higher yields in still water, but may be more likely to break (or be lost completely) in ocean locations that face significant waves. After planting, the farmers then tend their crops (remove debris, etc.). About 35 to 40 days later, the farmers harvest the seaweed, dry it, and then sell it to local buyers.

We became interested in this process for several key reasons. First, there had been a series of extension programs conducted in Indonesia, which provided advice to farmers on different aspects of farming, including optimal pod size, harvest timeframe, distance between pods and lines, etc. Few farmers followed the advice, and we were interested in understanding why. Second, this process is similar to that of other crop types, where the many different decisions over inputs and processes add up to determine yields. Third, seaweed farming has a relatively short crop cycle as

compared to many other crop types and there are several crop cycles each year. This made it interesting to study for two reasons: first, given the many cycles, there should be ample opportunities for farmers to learn through their own experimentation. Second, the short crop cycle allowed us to conduct a trial with farmers and to learn the results relatively quickly.

3.2 Experimental Design

From June 2007 to December 2007, we administered a survey to understand the demographics, farming practices, and learning practices of seaweed farmers in Indonesia (see Appendix Figure 1 for the project timeline). From a census of about 2706 farmers located in seven villages (24 hamlets), commissioned by us in 2006, we drew a random sample of five hundred farmers for the baseline survey, stratified by hamlet.[10] Out of these, 489 were located and participated in the baseline survey (see Appendix Figure 2 for the sample design). The baseline survey consisted of two parts: (1) a questionnaire that covered demographics, income, and farming methods, and (2) a "show and tell" where the enumerators visited one of the farmer's plots to measure and document his actual farming methods (see Section 3.3. for the data description).

From the list of farmers that participated in the baseline survey, we randomly selected 117 farmers to participate in an experimental trial to determine the optimal pod size for one of their plots (stratified by hamlet). The trials occurred between July 2007 – March 2008, shortly after the baseline was conducted for each farmer.[11] Farmers were told that the enumerators would vary the seaweed production methods across ten lines within one of their plots, with the farmer's assistance, and that they would be presented with the trial results afterwards.[12] All of the farmers that we approached participated in the trials and were compensated for doing so in two ways. First, we provided the necessary inputs for planting the lines, and guaranteed a given income from each line

[10]Once the initial 500 farmers were chosen, we showed the list to local village leaders. Those who had migrated prior to our survey or were no longer involved in seaweed farming were replaced.

[11]The exact timing of when the trial occurred varied from farmer to farmer given the timing of when a farmer had a plot free and also by the fact that our team could only manage a certain number of trials concurrently.

[12]This is consistent with the assumptions underlying Proposition 3, where farmers "know" they will receive a summary of the trial results, as well as with the practices of many extension services which demonstrate methods to farmers.

(that was calculated from our pilot tests), so that the farmers would at least break even. Second, we provided a small gift (worth $1) to each farmer to account for their time.

Farmers who participated in the trials were randomly assigned into one of two sub-treatments: *sort* (65 farmers) and *weight* (52 farmers). The sort sub-treatment was built around the idea that farmers had substantial variation in pod size even within their own plots (see Appendix Figure 3 for the distribution of sizes within farmers' own plots in this sub-treatment). Given this variation, we wanted to understand whether farmers could achieve a higher yield by systematically using a specific size within the support of those already used. Farmers were asked to cut pods as they usually would for the plot in question, and then the pods were sorted into small, medium and large groups. Working with the farmers, the enumerators attached the pods into the lines, by group (3 lines with small pods, 4 lines with medium pods, and 3 lines with large pods). The lines were then planted in the farmer's plot.

Despite the wide range of pod sizes used within a given farmer's plot, it is still possible that he could do better by moving to a size outside that range. The weight subtreatment was designed to address this issue by testing a broader set of initial pod sizes. To generate variation, the pod weights were initially set at 40g to 140g (in intervals of 20g) for the first few trials. However, in order to better reflect the ranges of weights used by the farmers, the weights were changed to 90g-180g for spinosim and 90g-210g for cottoni (both in intervals of 30g).[13] The pods of different sizes were randomly distributed across about 10 lines, with the enumerator recording the placement of each pod. The farmers were present for the trials and saw where each pod was planted on the lines.

In the weight subtreatment, we also tested whether farmers optimized distance between pods. We compared two distances, 15cm and 20cm, since the average distance between pods at baseline was around 15cm and past technical assistance programs had suggested larger spacing.

In both trial subtreatments, farmers were told to maintain the plots as they would any other plot. The enumerators returned to reweigh the seedlings twice while in the sea: once around day

[13]We ensured that the actual weights of the pods that were planted were within 10 grams of the target weight, and so it is best to think of these as bins within these averages. This adds some noise to the weight categories, biasing us against finding different effects by weight.

14, and again around day 28 (the exact dates depended on the conditions of the sea). On around day 35, the seaweed was harvested and weighed for a last time (once again, the date of harvest varied with conditions in the sea, farmer availability, etc.). Note that, in addition to being present for the planting, farmers were present for (and, in fact, helped with) the weighing, harvesting, and recording of results. Thus, they had access to all of the raw information generated by the trials.

In April-May, 2008, we conducted the first follow-up survey, which was designed to test whether farmers changed any of their methods after participating in a trial (see data section for details). These changes would have happened in the cycle after the trial: farmers would have had time to incorporate anything they learned on their own from the trial into the next cycle. This was conducted with a subsample of 232 farmers, which included all of the farmers who participated in the trials, as well as an additional set of farmers who were randomly selected from the baseline as a control group; 231 farmers completed the survey.

In May to June, 2008, we provided the results to the farmers. An enumerator gave each farmer a summary table that provided information on the returns from different methods and that highlighted which method had the highest yields.[14] The enumerators talked through the results with the farmers, describing the average weight of the pods that they typically used and the difference between that pod weight and the optimal pod weight (note that the optimal pod weight for sorting trials was the average of the optimal size conditional on the optimum being small, medium or large). For the weight trials, they were also told whether the optimal distance between pods was 15cm or 20cm. Appendix Figure 4 presents examples of the summary tables.

Approximately two months after the results were given to the farmers (July – August 2008), we conducted a second follow-up survey to determine if they changed their methods as a result of having received their trial results. Out of the original 232 farmers, 221 were found.

[14]We worked with a local NGO, who provided extension services to seaweed farmers and worked with us in conjunction for this survey, to design a simple, easy to understand table to summarize the trial results.

3.3 Data, Sample Statistics and Randomization Check

3.3.1 Data

Baseline Survey: This consisted of two parts: (1) a household questionnaire and (2) a "show and tell" where the enumerators visited one of the farmer's plots to document the actual farming methods.

The household questionnaire consisted of several modules. In the first section, we collected basic demographic information (household size, educational background, etc). In the second section, we collected data on how the farmer became involved in seaweed farming and how he obtained his plots, while we questioned the farmer to learn his beliefs on optimal farming methods in the third section. In the fourth section, we gathered data on each farmer's production function: labor costs, capital inputs, technologies employed, difference in methods based on seasonality and plot location, crop yields, etc. Next, we collected data on both learning and experimentation. The "learning" questions focused on external learning, such as: Where did the farmer gain his knowledge on production methods? Where does he go to seek new information, or to learn new techniques? The "experimentation" questions focused on internal learning/learning by doing, such as: Has the farmer experimented with different techniques? If yes, which types of techniques has he tried? Did he change his production methods all in one shot, or was it a step-wise process? Finally, we also collected data on income, expenditures, assets and loans.

After the questionnaire was complete, the enumerator conducted the "show and tell" to document the farmer's actual production methods.[15] This typically took about 15 minutes to two hours, depending on how difficult it was to access a plot. During the show and tell, we collected information on the type of line used, the sizes of a random sample of pods, the distance between a random sample of seedlings, the distance between seaweed lines, and so forth.

[15]For some farmers, the show and tell was conducted on another day, as a low tide was neccesssary in order to access the plots.

Experimental Trial Results: We compiled data from each of the experimental trials. Specifically, we documented each plot location, where each pod was planted within a plot, and the weight of each pod during each of the three weighings. As such, we can compute the yield and the return from each pod that was planted.

Follow-up Surveys: In both follow-up surveys (one after the trial and one after the summary data was seen), we collected information on farmers' self-reported changes in farming techniques. Additionally, we again conducted the "show and tell" module to determine if the treatments affected measured methods.

3.3.2 Baseline Sample Statistics and Randomization Check

As Panel A of Table 1 illustrates, most farmers (83 percent) were literate. Most farmers had been farming for about 18 years, with about half reporting that they learned how to farm from their parents (Panel B). As is typical in the agricultural sectors in many developing countries, a large share of farmers (about a quarter) had a loan from the buyer to whom they sell their crop. About a third of the sample grew the cottoni strand at baseline. The "show and tell" revealed that the mean pod size was about 105 grams, while both the distance between pods and between lines was about 15 cm.

In Appendix Table 1, we provide a randomization check across the control and both subtreatment groups. We test for differences across the groups on ten variables at the baseline, where these variables measure demographic information and key farming characteristics. Columns 1 - 3 provide the means for the control, sort and weight groups respectively, while Columns 4 - 6 provide estimates of the differences in means across groups. As illustrated in Columns 4 through 6, out of 30 comparisons we consider, 3 are significant at the 10 percent level (consistent with what would be expected by random chance). We additionally compute the p-value from a test of joint significance across all 10 variables; this is reported in the final row in Columns 4 - 6. We fail to reject the null that all the differences are equal to zero for each of the three comparisons.

4 Results

4.1 Tests of the Model

4.1.1 Baseline Results

A basic implication of learning through noticing is that some farmers will not keep track of input dimensions that influence returns. Table 2 presents baseline sample statistics for the 489 farmers that were included in this survey. Panels A and B document self-reported beliefs on current methods and optimal methods. For these questions, we present the percentage of farmers who were unable to provide an answer in Column 1, and then provide the means, standard deviations, and number of observations for those who provided an answer in Columns 2, 3, and 4. In addition, Panel C provides information on whether farmers had previously actively changed their production practices along a number of dimensions.

Result 1: Farmers only attend to certain dimensions of the production function.

A vast majority of farmers were inattentive to some input dimensions, particularly pod size. Eighty six percent could not provide an answer for their current pod size at baseline (Column 1 of Panel A), while 87 percent of farmers did not even want to hazard a guess about what the optimal pod size should be (Panel B).[16] Since many farmers fail to notice key facts about their pod sizes, it is perhaps unsurprising that a broad range of sizes is observed both across (see Figure 1) and within (see Appendix Figure 3) farmers, and that only 12 percent indicated they had previously tried different sizes (Panel C).

On the other hand, farmers appear to have been more attentive to other input dimensions. For example, they appeared attentive to both the distance between knots that secure the pods to a line and the distance between lines. Unlike pod size, very few farmers (about 1 to 2 percent) could not provide an answer for their distance between lines (Panel A). In addition, most farmers had

[16] In fact, the enumerators commented to us that many farmers could not answer these questions, even when probed further.

an opinion about the optimal distance between both the knots and lines (Panel B). Given that the farmers are relatively more attentive to the distance between knots and lines, we might expect that the actual distances employed would exhibit less variance than the pod sizes. Using the means and standard deviations from Table 1 (Panel B), we can compute the coefficient of variation to compare the relative variances across inputs: the coefficient of variation for the distance between lines (0.13) and pods (0.10) is much smaller than that for pod size (0.27), indicating that the practices for these inputs are relatively less variable across farmers. Finally, note that an extremely small percentage of farmers reported trying out different distances between lines and pods in the past (Table 2, Panel C), suggesting that social and public learning may contribute to farmers' beliefs about the optimal distance between pods and lines.

Why do farmers notice distance, but not pod size? It is hard for us to conclusively say, but the model suggests some reasons why this may be the case. The model implies that farmers are less likely to attend to a relationship when the option value is lower or when the cost of paying attention is higher, both of which plausibly hold when comparing pod size to distance between pods. For example, our empirical work below suggests there is more heterogeneity in the optimal pod size than in the optimal distance across farmers, implying a lower value of attending to data from neighbors' plots. While this feature may slow down learning in a traditional learning model, in our model it can completely shut it down. In terms of the costs of attending, a farmer who notices variation in yields at the end of a crop cycle can look back at whether distance matters since the distance between knots stays the same through the crop cycle. However, to learn the importance of pod size, he needs to keep track of the initial size at the time of planting. Independent of the precise reason, the rest of the empirical analysis proceeds from the idea that farmers were more likely to attend to distance than to pod size.

Result 2: Farmers fail to optimize along unattended-to dimensions.

A second implication of the model is that when farmers do not attend to some input dimension, their choices along that dimension can be far from optimal. For the 117 farmers that participated

in the experimental trials, we have data on optimal pod size. Table 3 summarizes the predicted percentage income gain from moving away from initial pod sizes.[17] Panel A reports this estimate for farmers in the sort sub-treatment, providing information on both the predicted gain a farmer could achieve by changing the size of each pod from his baseline average to the best performing size and the predicted gain by changing the size of each pod from the worst performing size to the best among sizes used at baseline. Panel B then presents the predicted income gain a farmer could achieve by changing the size of each pod from his baseline average to the best performing size in the weight sub-treatment. We provide the median gain for farmers in Column 1, and provide the confidence interval of the estimate in Column 2.[18] Note that, on average, each trial had around 300 pods per farmer so we have a reasonable number of observations with which to calculate these returns. Additionally, we estimated a regression of yield on size dummies for each farmer separately, and passed an F-test of joint significance in every case.

On net, the results indicate that farmers are potentially forgoing large income gains by not noticing and optimizing pod size.[19] In the sort sub-treatment, the median estimated percentage gain in income by moving from the average to the best performing size is 7.06, while the median estimated gain by moving from the worst to the best is 23.3 (Panel A). In the weight sub-treatment, the estimated gain by moving from the average at baseline to the best size is 37.87 (Panel B). Given the wide heterogeneity in returns, many individual farmers could even potentially increase their incomes by much more than what is typical across the farmers: the data suggests that many farmers are simply losing out on potential returns.[20] Most strikingly, the gains from the sort sub-treatment suggest that farmers would have done much better by systematically using a specific size

[17]Note that to compute these predicted changes to income, we make several strong assumptions. First, we assume that past seaweed prices are consistent with the future ones; this may be unrealistic as the price may fall if all farmers increase their yields. Second, we assume that farmers do not change other methods (have fewer cycles, harvest earlier, etc.) if their yields change. Thus, this evidence should be viewed more as *suggestive*, rather than as the *counterfactual* of the change in income if farmers changed their methods.

[18]Given the wide heterogeneity in the results, the median is likely a better measure of what is typical than the mean.

[19]Note that in the sort sub-treatment, about half the farmers were told that their most productive bin was their largest bin, while about 30 percent were told they would be most productive by moving to the smallest bin. Similarly, in the weight treatment, about 55 percent were told that they should increase their initial sizes.

[20]Since we use the data to pick the optimal size for each farmer, note that the gains from switching estimate is biased upwards. However, by the strong law of large numbers it is an asymptotically consistent estimator (as the number of pods per size per farmer gets large).

within the support of those already used. As discussed in greater detail below, this fact indicates that it is unlikely that farmers' failure to optimize purely reflects a failure of experimentation.[21]

Farmers appear to perform better in setting their distance between pods – a dimension they seemed to notice at baseline. Results from the weight sub-treatment indicate that, for 80 percent of farmers, the optimal distance between pods was 15cm. Given that most farmers were at 15cm to begin with, these data suggest that very few farmers would do better by changing to 20cm.[22]

Taken together, these findings suggest that many farmers failed to notice pod size and were not optimizing this size, while many farmers noticed distance between pods and were optimizing this distance (at least within the support of distances tested in the trial). The results suggest that inattention contributes to a failure to optimize and hinders learning by doing. We now analyze the farmers' response to the experimental trial, which provides a stronger test of the theory.

4.1.2 Response to the Trial

The model predicts how farmers should respond to the experimental trial. Specifically, we would expect that participating in the trial by itself may not have large effects on future behavior if farmers tend not to notice the dimensions they fail to optimize. However, farmers should be more likely to respond when they are presented with a summary of the trial findings, as the summary is easier to process.[23] When farmers respond, they should bring their practice more in line with what per-

[21] However, comparing the gains-from-switching estimates across sort and weight sub-treatments indicates that farmers may have done even better by moving to a size outside of the support of those already used, which can be interpreted as suggesting that a lack of experimentation also contributes to a failure to optimize.

[22] Given that there appears to be substantial heterogeneity across farmers in the optimal size (see Appendix Figure 5), this fact is also suggestive that there is more heterogeneity across farmers in the optimal size than the optimal distance between pods.

[23] The prediction that farmers will respond more to the trial plus the summary by more than to the trial by itself relies in part on an assumption that simply being asked to participate in the trial does not significantly draw farmers' attention to pod size – i.e., being asked to participate does not lead farmers to significantly update the probability they place on pod size mattering, π. While similar assumptions may not hold in other contexts (see, e.g., Zwane et al. 2011), several pieces of data help build the case for this assumption in this context. As mentioned above, few farmers in other parts of Indonesia previously took up NGO advice on pod size. Indeed, this was one of the reasons we became interested in running the current experiment in the first place. In fact, very few farmers at baseline (roughly 10 percent) indicated that they would change their farming methods simply in response to an NGO or government recommendation, or in response to advice from a friend (Appendix Table 2), while far more farmers (roughly 40 percent) indicated that they would change their practice in response to results from other plots. These results suggest a hesitation among these farmers to take advice at face value.

formed well at the trial. If farmers are further from the optimum for some inputs (such as pod size) than others (such as distance between pods), we would expect that learning the trial results should have larger effects on the inputs that are further from the optimum.

We begin by exploring the effect of participating in a trial on production outcomes. Specifically, for each farmer i in hamlet v, we estimate the following model:

$$Y_{ivt} = \beta_0 + \beta_1 F1_t + \beta_2 F2_t + \beta_3 Trial_{iv} + \beta_4 Trial_{iv} * F1_t + \beta_5 Trial_{iv} * F2_t + \alpha_v + \mu_{ivt} \quad (10)$$

where Y_{ivt} is the production choice at time t, $F1_t$ is an indicator variable that denotes that the observation comes from the first follow-up after the experimental trial, $F2_t$ is an indicator variable that denotes that the observation comes from the second follow-up after the summary findings were presented to farmers, and $Trial_{iv}$ is an indicator variable that denotes trial participation. We also include a hamlet fixed effect, α_v, as the randomization was stratified along this dimension.[24] There are two key parameters of interest: β_4 provides the effect of participating in the trial prior to obtaining a summary of the trial results, while β_5 provides the effect after these results are provided.

Table 4 presents these results. In Columns 1 and 2, the outcome of interest is the self-reported measure of whether the farmer has made any changes in their production techniques.[25] In Column 1, we report the coefficient estimates from Equation (10); in Column 2, we report the estimates from a model that additionally includes farmer fixed effects. Columns 3 and 4 replicate the analysis in the first two columns, but with the enumerator measured pod size as the outcome variable. We estimate all models using OLS and all standard errors are clustered by farmer.[26]

[24]Note that the inclusion of a hamlet fixed effect does not significantly influence the results; results from specifications that omit this fixed effect are available upon request.

[25]This includes pod size, distance between lines and distance between pods.

[26]The findings are similiar if we cluster by hamlet instead. The findings for whether farmers changed their farming techniques are also similar if we estimate the equation as a probit rather than OLS.

Results 3 and 4: Participating in the trial by itself had little effect on farmers' decisions. Presenting farmers with a summary of the trial's findings was more effective.

Consistent with the model's predictions, just participating in the trial had no effect on the subsequent production outcomes, but learning the results led to large changes in production methods.[27]We do not observe a significant change in self-reported changes to farming techniques from participating in the trial, prior to when farmers received the summarized results (Table 4, Column 1). However, about 16 percent more farmers reported changing a technique after receiving the results, which is about one and a half times the mean of the dependent variable (Column 1). Adding farmer fixed effects does not significantly alter the estimated coefficient (Column 2). It is possible that some of the results from this self-reported measure may be driven by farmers wanting to please the enumerators after participating in the trial, though this is unlikely as the control group had also been receiving visits from enumerators to both survey and measure their farming practices. Nonetheless, we turn to the enumerator measured results, which are less likely to suffer from this type of bias.

We do not observe a significant change in pod size from participating in the trial, prior to when farmers received the summarized results (Columns 3 and 4). After receiving the summary, however, pod size increases by about 7 grams (on average) for those those in the treatment group, relative to the control. This is significant at the 10 percent level in the basic specification (Column 3) and the positive sign of the effect is consistent with the average recommended change in pod size from the trials. However, while the coefficient estimate remains roughly the same (7.3 grams) when including farmer fixed effects, the significance level falls below conventional levels of significance (p-value = 0.14) due to an increase in the standard error (Column 4). Thus, while farmers did not appear to consider pod size to be an important part of the production process prior to the trials, providing summary information on their optimal pod size appeared to change their use of this

[27]We only varied distance between pods in the weight sub-treatment. In the sort sub-treatment, we kept all aspects of the farmer's production methods constant, except to sort the pods by their natural variation. As a result, we only present the distance results in Table 5 where we disaggregate results by the sub-treatments.

production input.[28]

We next explore the sort and weight sub-treatments. Here, we provide coefficient estimates when we include dummy variables for the sort and weight treatments as controls (S_{iv} and W_{iv}), as well as the interaction of these variables with the indicators for follow-up status:

$$Y_{ivt} = \beta_0 + \beta_1 F1_t + \beta_2 F2_t + \beta_3 S_{iv} + \beta_4 W_{iv} + \beta_5 S_{iv} * F1_t + \beta_6 S_{iv} * F2_t \quad (11)$$
$$+\beta_7 W_{iv} * F1_t + \beta_8 W_{iv} * F2_t + \alpha_v + \mu_{ivt}.$$

β_5 through β_8 are the coefficients of interest, providing the effect of the respective sub-treatment in the respective time period. The first two columns report the coefficient estimates where self-reported change is the outcome of interest, while the next two columns report the specifications for enumerator-measured pod size. Distance between pods is the outcome variable in Columns 5 and 6 (note that this was only tested in the weight sub-treatment). We present results with (odd columns) and without (even columns) individual fixed effects; standard errors are clustered by farmer.

The findings are consistent with the predictions of the model. Farmers do not appear to make significant changes in production methods simply from participating in either type of trial (Column 1 and Column 2). However, in both types of trials, they report making large and significant changes in their production techniques after they are presented with the results. Likewise, we find no effect of merely participating in the trial on enumerator measured pod size (Columns 3 and 4), but observe positive effects of being presented with the trial's results (while the effect in the weight

[28]In Appendix Table 3, we disaggregate the results by whether farmers were told to increase or decrease pod size in the follow-ups. To do so, we interact the interaction of the treatment and follow-up variables with an indicator variable that denotes whether the farmer should increase their pod size (Increase), controlling for the main effect of the recommendation (as well as all double interactions between the recommendation and the treatment variable and the recommendation and the follow-up variable). Once again, we present results with (Column 1) and without (Column 2) individual fixed effects; all standard errors are clustered at the farmer level. We observe that farmers who were told to increase pod size did so both after the first and second follow-up. However, the recommendation itself is an endogenous variable; thus, it is possible that this is simply capturing the fact that if farmers are randomizing with respect to pod size, those who "should go bigger" are those who had abnormally low draws of pod size at baseline. Thus, in expectation, those farmers would in fact go bigger the next period even if they continue to randomize. Therefore, these results are difficult to interpret.

sub-treatment goes in the right direction, the effect is only significant in the sort sub-treatment).

Finally, we find no effect of the weight treatment on the enumerator measured distance between pods in either period. This is consistent with the model: Unlike pod size, farmers appeared to previously notice distance, had beliefs on the optimal distance, and tended to be at the optimum (at least within the support of distances tested in the trial). As a result, we would not expect large changes in distance as a result of either participating in the trial or receiving its results. However, it is important to note that while this result is consistent with the model, the insignificant result on distance could also be driven by the smaller sample size.

4.2 Other Interpretations

While the experimental findings closely track the theory, other explanations of these findings are certainly possible. We discuss these alternatives in this Section.

First, perhaps we have missed important costs and, in fact, farmers' pod size strategy is optimal. For example, perhaps carefully cutting pods to match a given pod size is costly in terms of labor effort, and, even though yield may be higher, utility would not be. The challenge for this explanation, however, is the finding that farmers change their behavior after they were shown the summary results. If farmers were at the optimum, there would be no reason for them to change their behavior. This change instead suggests that the farmers themselves felt they were not at the optimum based on the data they saw.

Second, one might argue that the failure to learn comes from a failure to experiment. While farmers may be underexperimenting, this alone cannot explain our full set of findings. In our study, farmers participate in a trial on their own plots. They see the data from these experiments (or rather, in our model, they have access to the data), yet this does not affect their behavior. Even more problematic for this explanation, farmers are *constantly* experimenting: by not paying attention to pod size they are generating variation in pod size. In fact, we see from the sort treatment that even within the range of pod sizes that they normally use, they are not picking the highest yield one.

A third explanation is that maybe the farmers noticed pod size, but are simply innumerate.

They could not learn from participating in the trial – or even from the natural variation in the data – because they could not do the math necessary to calculate the return. Note, however, that literacy rates tend to be high, with about 83 percent of the farmers literate in the baseline. Additionally, the fact that farmers noticed and were able to optimize on distance suggests that numeracy may not be the main problem or it would pose a problem along this dimension as well.

Perhaps the biggest stumbling block for all these (and other) explanations is the link between failures to learn and self-reported knowledge. Our theory makes a clear prediction: failures to learn will be centered on dimensions that people lack knowledge of, as in the case of pod size. While other explanations may explain why people may fail to learn (or even fail to learn from experiments they participate in), it is less clear why those failures would be linked to what is noticed. Limited attention appears to be the most plausible explanation of this key fact. In addition to our model, other models of inattention could produce similar effects (e.g., Sims 2003 or Gabaix 2011). As typically written, these other models would have trouble explaining why farmers did not attend to pod size when it is so important, since such models typically combine inattention with a sort of rational expectations assumption. This feature is not inherent to the precise way they formalize inattention, however. In principle, it seems possible to combine models such as theirs with a model of learning the production function in a way that would generate results that are similar to ours. What these data do single out is a model that links what is noticed to what is learned.

5 Model Extensions and Further Applications

This section presents extensions to the basic model that are important in settings outside of the main empirical application, and discusses further applications.

5.1 Adopting a Technology

So far, we have taken the technology adoption decision as given. Now suppose that in addition to deciding how to use a technology, the farmer also decides *whether* to use it. Extend the learning

by doing model so that, each period, the farmer faces an additional decision of whether to farm at all. If he does not farm, he receives payoff $\bar{u}$, normalized to equal 0.

To limit the number of cases considered, suppose that $\mu > 0$, which implies that the farmer will farm in the first period. Recalling that s_0 denotes the input level that maximizes the true production function, $f_0(s)$, further suppose that $\max\left\{f_0(s_0) - m, \frac{1}{S}\sum_{s'} f_0(s')\right\} = f_0(s_0) - m > 0$, so a farmer that employs the best practices is better off farming than not given the underlying technology, and using a particular input level rather than not measuring.

Whether a selectively attentive farmer decides to adopt a profitable technology depends on whether he attends to the proper inputs. We will say that, given what a farmer selectively attends to, a technology is *filter incongruent* if it would no longer be profitable if the farmer only optimizes with respect to attended to inputs and randomizes with respect to the other inputs. Formally:

Definition 5.1. Fix whether the farmer attends to the input level in the initial period. The technology is *filter incongruent* if the farmer does not attend and $\frac{1}{S}\sum_{s'} f_0(s') < 0$.

Intuitively, the farmer is less likely to adopt a filter incongruent technology. While these technologies would be profitable if adopted correctly, the farmer will not learn to optimize them because he fails to attend to important input dimensions and thus believes that they are unprofitable. To formalize this intuition, we will assume away tangential technical issues by simplifying the farmer's space of potential strategies as well as the environment.

Assumption 5.1 (Symmetry)**.** The farmer is restricted to using strategies that are symmetric across pieces of land in a given period. For example, if he does not measure or attend to the input level applied to piece of land p' at $t = 1$, then he does not measure or attend to the level applied to piece of land p'' at $t = 1$.

Assumption 5.2 (Deterministic environment)**.** If the farmer does not measure the input level applied to any piece of land, then $N/S = n$ pieces of land are randomly assigned to input level $s = 1$, n to $s = 2$, etc. Further, $\sigma_\varepsilon^2 = 0$.

Assumption 5.1 simplifies matters by helping to guarantee that the farmer finds it optimal not to measure particular input levels when indifferent between them. Assumption 5.2 is more important and greatly simplifies the presentation and proofs by implying that the farmer's beliefs evolve deterministically.

Proposition 4. *Suppose Assumptions 5.1 and 5.2 hold.*

1. *If the technology is filter congruent, then the farmer continues to farm in the second period.*
2. *If the technology is filter incongruent, then the farmer does not continue to farm in the second period.*

Proposition 4 indicates that farmers are less likely to adopt filter incongruent technologies. Assuming their priors come from experience with other technologies, this suggests that they are less likely to adopt technologies when the relevant inputs differ from those of technologies they have experience with. This provides one reason why farmers may be quicker to adopt technologies that are "compatible" or "congruent" with current practices (Rogers and Shoemaker 1971, Brandner and Strauss 1959, Tornatzky and Klein 1982, Fliegel and Kivlin 1966). For example, Brandner and Strauss (1959) argue that farmers who previously adopted hybrid corn were more likely to subsquently adopt hybrid sorghum, which could result from the common practices involved in optimally planting hybrid varieties. An even more unique implication is that experience with related technologies may actually be harmful. In the skillet example, experience using non-stick pans may have "taught" you that the precise details of cleaning the pan do not matter and, therefore, such "knowledge" may keep you from learning not to use soap. A less experienced cook may actually be in a better position to notice this crucial detail.

5.2 Learning from a Best-Practice Demonstrator

Suppose now that there is an additional period 0, where the farmer does not farm his own plot, but rather observes a demonstration by an individual with knowledge of the underlying production function. We will call this individual a *best-practice demonstrator.* The best practice demonstrator

optimally farms the farmer's plot. Since the farmer is selectively attentive, he may not attend to the demonstrator's choice of input s_0 and he must thus hold some beliefs over what the demonstrator does. We assume that, whatever these beliefs are, they reflect ignorance in the sense that if he does not attend to what the demonstrator does, his posterior remains symmetric across input levels. That is, the farmer has no reason to suspect that any one piece of land is farmed differently than another, or that any one input level is more likely to be selected than another. This assumption is likely to be appropriate so long as the demonstrator does not communicate his input choice to the farmer. (More on this below.)

Again, whether the farmer profitably uses the technology depends on whether it is filter congruent:

Proposition 5. *Consider a farmer who learns from a best-practice demonstrator in period* 0 *and then decides whether and how to farm in periods* 1 *and* 2*. Under Assumptions 5.1 and 5.2:*

1. *If the technology is filter congruent and the farmer farms in period* 1*, then he continues to farm in period* 2*.*

2. *If the technology is filter incongruent and the farmer farms but does not attend in period* 1*, then he does not farm in period* 2*.*

Remark 5.1. The proof of Proposition 5 makes clear that if the farmer attends to what the demonstrator does, he necessarily chooses to farm in every period because the expected profitability of farming is then bounded below by $N \cdot (f(s_0) - m) > 0$ given Assumption 5.2.[29] In particular, our model suggests that behaviors deviating from persistent adoption following a successful demonstration trial signal a failure to notice.

Proposition 5 indicates that inattentive agents can experience disappointment. For example, farmers can observe data–perhaps in a demonstration–that leads them to believe that they can use a technology profitably, but then experience low returns when they try to use it by themselves. Some

[29]The qualifier "if the farmer farms in period 1" thus handles the case where the farmer does not attend to what the demonstrator does in period 0.

prior evidence can be interpreted in this light. Prominently, the fertilizer experiments by Duflo, Kremer and Robinson (2008) study the impact of being invited to observe a trial on a farmer's plot and of having a trial performed on one's own plot. In both cases, they find that a modest percentage of farmers adopt fertilizer the next year. Most interesting for our purposes, though, is what happens next. Adoption rates decline in subsequent years, implying that farmers could have disappointed with the results when they tried it on their own. In traditional models, participation in trials or observing trials is sufficient to encourage adoption. In our model, the impact of this type of teaching activity depends on whether the technologies are "filter congruent"; i.e., how well they line up with the practices agents notice.

5.2.1 Adding Communication

Note that an additional form of communication can help when the technology is filter incongruent: the best-practice demonstrator can explain his strategy to the farmer either before or after the demonstration. When this strategy involves the use of a particular size, s_0, then the farmer would effectively have access to the full history h when updating his beliefs (so long as he processes the communication). As a result, this farmer is more likely to profitably adopt the technology.

In practice, the demonstrator makes many input choices and communicating these choices to the farmer is costly (Niehaus 2011). He should only communicate choices along dimensions the farmer is predisposed not to attend to and are relevant for the task at hand. The effectiveness of communication is then increasing in the degree to which the demonstrator has knowledge of the farmer's mental model; i.e., of what he does and does not attend to. This discussion speaks to the difficulty of agricultural extension (Evenson 2001, Gautam 2000, Anderson and Feder 2004). Simply sending an extension worker will not guarantee subsequent adoption and proper use of a profitable technology, even if the worker demonstrates how to use it. For certain technologies, effective communication is key.

6 Conclusion and Broader Discussion

In this paper, we propose that learning through noticing influences technology adoption and use. Beyond the specific predictions we discuss in the paper, this perspective also forces us to re-think the role of experience and education in promoting efficient use of technologies.

Learning through noticing alters the standard intuition that experience guarantees effective technology use (see, for example, Nelson and Phelps 1966, Schultz 1975, Foster and Rosenzweig 2010). While experience will lead people to optimize along dimensions they attend to, our framework and empirical findings suggest that they may be far from optimizing along the dimensions that they fail to notice in the first place.[30] This insight may help us understand persistent differences in practices across agents. For example, a large literature posits that persistent differences across farmers must reflect differences in endowments or constraints (Griliches 1957, Schultz 1963, Gerhart 1975, Croppenstedt et al. 2003). In a recent paper, Suri (2011) argues that heterogeneity in adoption of hybrid maize in Kenya can be explained by heterogeneity in the benefits and costs of adoption, which farmers know from experience. We offer an alternative explanation: farmers may act differently because of differences in what they notice. Quite simply, some farmers may not notice the very features of the technology that make it profitable.[31]

Similarly, our analysis has important implications for education. First, it suggests that training programs can be useful, not only for new technologies, but also for existing technologies that individuals may have had experience with. To understand when this is the case, we should look at indicators not only of incorrect beliefs but also of insufficient attention. Moreover, our find-

[30] While bandit models (Gittins 1979) and "local" models of learning by doing (Conley and Udry 2010) also allow for a persitent failure to optimize, the predictions of such models are very different. In such models, a lack of data explains the learning failures. To diagnose a problem, we should look for indicators that people do not have sufficient data available, perhaps due to insufficient experimentation. To solve these problems, we need to provide the relevant data, or to subsidize their collection. When the binding constraint is instead a failure to notice, looking for instances where people do not have data available is not sufficient to diagnose problems; providing this data is not enough to solve them.

[31] To test this hypothesis, it would be necessary to collect data on what people notice, extending our approach from the context of seaweed farming. For example, in the context of maize production in Kenya, it would be useful to understand whether persistent hybrid adopters are more likely to be able to answer questions about what they do along dimensions important when using hybrids, e.g., about when and how they apply fertilizer or exactly how much fertilizer they use (Duflo et al. 2008). Importantly in this context, even nonadopters are likely to have used hybrid maize at some point in the past (Suri 2011), and so in principle should be able to answer these questions.

ings also inform the design of these programs: simply demonstrating how to use a technology or even listing out dimensions of a technology – many of which the selectively attentive agent may have already chosen to ignore – may not lead to optimal use. For example, this point may help explain the difficulties (and relatively low success rates) of agricultural extension, which often provide a lot of information to farmers in a short amount of time (Evenson 1997, Gautam 1999, Anderson and Feder 2004). Rather, training may be more powerful when it focuses on the aspects that agents typically fail to notice, and highlights best practices along those dimensions. As such, future research on training and extension programs should help disentangle how the presentation and framing of information affects technology use and profitability. Second, our findings speak to the fact that *general education* can also be beneficial through increasing agents' ability to process information along multiple dimensions (through lowering e). For example, Ceci (1991) and Neisser et al. (1996) provide evidence of a causal link from schooling to general cognitive skills like working memory, which provides a microfoundation for the widely held belief that schooling can increase the speed of learning and the ability to adapt to new circumstances (Nelson and Phelps 1966, Schultz 1975, Rosenzweig 1995, Foster and Rosenzweig 2010).

In general, this discussion suggests a more nuanced view of human capital. Simply being able to encode information along more dimensions – a very basic cognitive capacity – may have effects on learning capacity across many problems. More sharply, human capital is no longer summarized by exposure to data or knowledge of facts: Embodied in individuals is information about what to notice and what to neglect.

Table 1: Baseline Demographic Characteristics and Farming Practices

	Mean (1)	Standard Deviation (2)	N (3)
Panel A: Demographic Characteristics			
Monthly Per Capita Expenditures	369543	348368	487
Age of HH Head	43.08	11.87	474
Number of Assets	8.09	3.23	487
HH Head is Literate	0.83	0.38	480
Panel B: Seaweed Farming Practices			
Years Farming Seaweed	18.36	7.15	475
Learned to Farm Seaweed from Parents	0.50	0.50	487
Has a Loan from Person to Whom Sells Seaweed	0.28	0.45	353
Number of Days in Previous Cycle	36.74	7.75	487
Mean Distance Between Lines at Baseline (Enumerator Measured)	15.47	1.96	486
Mean Distance Between Pods at Baseline (Enumerator Measured)	15.20	1.47	486
Mean Pod Size at Baseline (Enumerator Measured)	105.74	28.72	487
Cottoni Strand	0.34	0.47	487

Notes: This table provides sample statistics on the farmers' demographic characteristics and seaweed farming practices from the baseline survey.

Table 2: Baseline Survey Responses on Process and Practices

	Percent Unable to Provide Answer (1)	Mean (2)	Standard Deviation (3)	N (4)
Panel A: Self-Reported Current Production Methods				
Current Pod Size	86%	118.11	57.01	70
Length of Typical Line	2%	5.05	1.04	481
Distance Between Lines	1%	16.49	3.14	482
Panel B: Beliefs on Optimal Production Methods				
Optimal Pod Size	87%	148.26	248.45	63
Optimal Distance Between Knots	2%	15.97	2.84	481
Optimal Distance Between Lines	2%	16.39	3.01	481
Optimal Cycle Length	1%	37.43	7.14	486
Panel C: Self-Reported Changes in:				
Pod Size		0.12	0.32	446
Distance Between Lines		0.01	0.08	484
Cycle Length		0.09	0.28	475
Distance Between Knots		0.04	0.21	478
Strain		0.51	0.50	487
Any Aspect of Farming, Other than Change in Strain		0.19	0.39	487
Any Aspect of Farming		0.57	0.50	487

Notes: This table provides sample statistics on farmers' responses from the baseline survey.

Table 3: Estimated Percent Income Gain from Switching to Trial Recommendations

	Median Gain (1)	95 Percent Confidence Interval (2)
Panel A: Sort Treatment Group		
Gain from Moving from Average to Recommendation	7.06	[2.92 ,14.19]
Gain from Switching from Worst Bin to Best Bin	23.3	[19.00, 28.18]
Panel B: Weight Treatment Group		
Gain from Moving from Current Pod Size to Best	37.87	[23.60, 58.86]

Table 4: Effect of Participating in the Trial on Self-Reported Techniques and Measured Pod Size

	Changed Farming Techniques		Pod Size	
	(1)	(2)	(3)	(4)
Trial Participation	-0.084		-2.184	
	(0.051)		(3.610)	
After Trial	-0.146	-0.148	-11.333	-11.661
	(0.048)***	(0.057)**	(3.003)***	(3.578)***
After Summary Data	-0.145	-0.150	-13.587	-13.859
	(0.050)***	(0.061)**	(2.896)***	(3.496)***
Trial Participation * After Trial	0.072	0.079	-2.051	-1.550
	(0.060)	(0.071)	(4.411)	(5.306)
Trial Participation * After Summary Data	0.162	0.171	6.951	7.316
	(0.069)**	(0.084)**	(4.095)*	(4.982)
Hamlet Fixed Effects	X		X	
Farmer Fixed Effects		X		X
Observations	684	684	684	684
Mean of Dependent Variable for the Control Group:				
After Trial	0.10	0.10	97.68	97.68
After Summary Data	0.11	0.11	95.39	95.39

Notes: This table provides the coefficient estimates of the effect of treatment on farming methods after the trial (follow-up 1) and after observing the summary data (follow-up 2), conditional on baseline farming methods. The trial participation dummy indicates that the farmer belongs in either the sort or weight treatment group. Changed Farming Techniques includes self-reported changes in pod size and distances, while pod size is enumerator measured. All regressions are estimated using OLS and standard errors are clustered at the farmer level. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Effect of Participation in Sort Versus Weight Trials

	Changed Farming Techniques		Pod Size		Distance Between Pods	
	(1)	(2)	(3)	(4)	(5)	(6)
Sort * After Trial	0.089	0.100	3.944	4.657		
	(0.065)	(0.077)	(4.461)	(5.310)		
Weight * After Trial	0.052	0.053	-9.257	-8.929	0.289	0.304
	(0.075)	(0.089)	(6.610)	(7.882)	(0.328)	(0.387)
Sort * After Summary Data	0.141	0.153	10.908	11.768		
	(0.075)*	(0.091)*	(4.418)**	(5.286)**		
Weight * After Summary Data	0.187	0.192	2.185	2.093	0.226	0.172
	(0.095)*	(0.114)*	(5.819)	(7.002)	(0.303)	(0.362)
Hamlet Fixed Effects	X		X		X	
Farmer Fixed Effects		X		X		X
Observations	684	684	684	684	499	499
Mean of Dependent Variable for the Control Group:						
After Trial	0.10	0.10	97.68	97.68	15.39	15.39
After Summary Data	0.11	0.11	95.39	95.39	15.27	15.27

Notes: This table provides the coefficient estimates of the effect of the different treatments on farming methods after the trial (follow-up 1) and after observing the summary data (follow-up 2), conditional on baseline farming methods. Changed Farming Techniques includes self-reported changes in pod size and distances, while pod size and distance between pods are enumerator measured. All regressions are estimated using OLS and standard errors are clustered at the farmer level. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Figure 1A: Baseline Pod Sizes

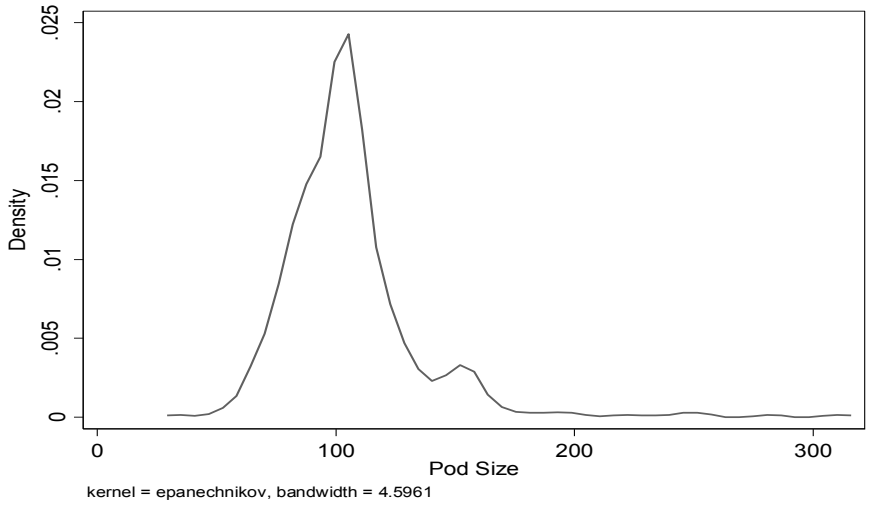


Figure 1B: Baseline Pod Sizes for Cottoni Growers

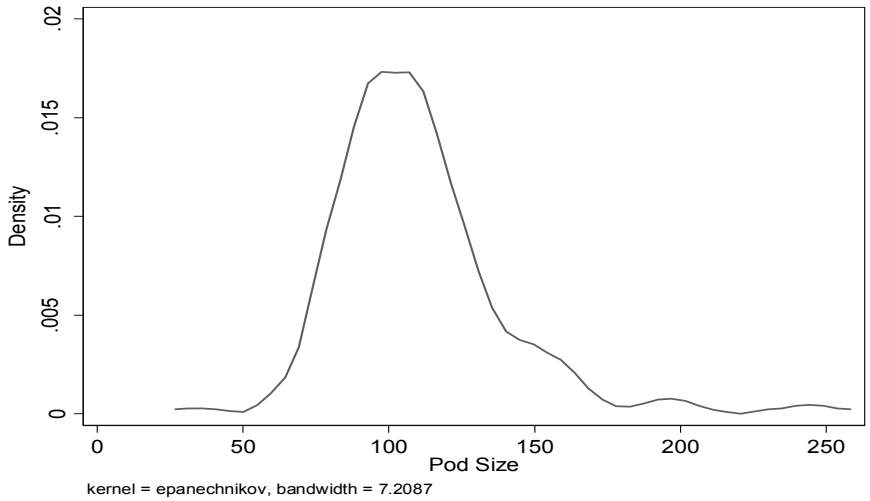


Figure 1C: Baseline Pod Sizes for Spinosim Growers

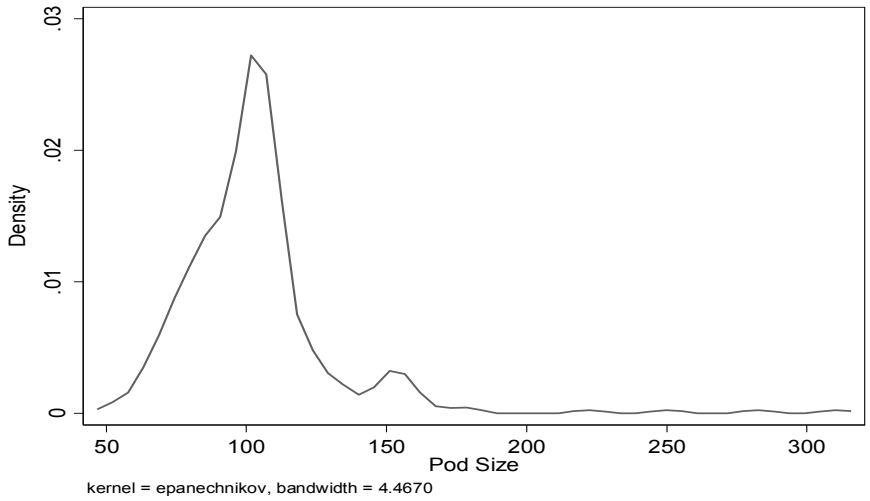
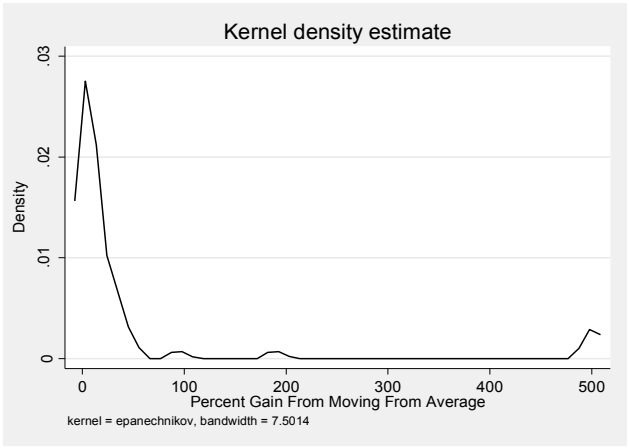
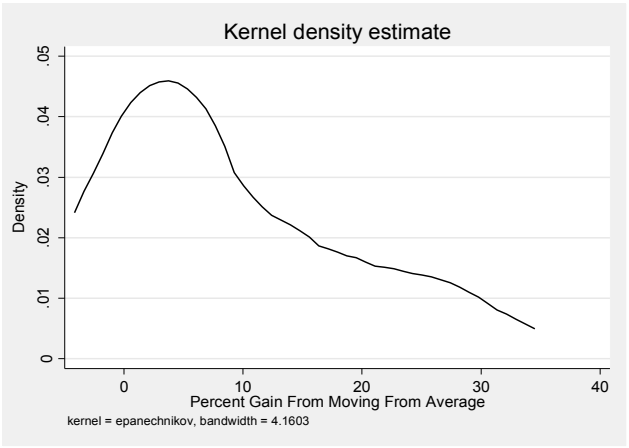


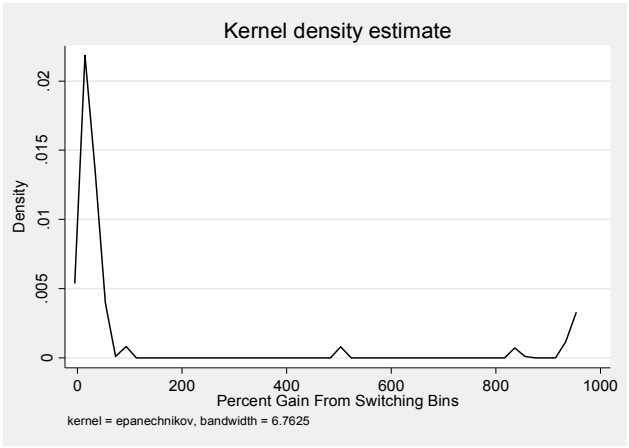
Figure 2: Distribution of the Estimated Percent Income Gain from Switching to Trial Recommendations

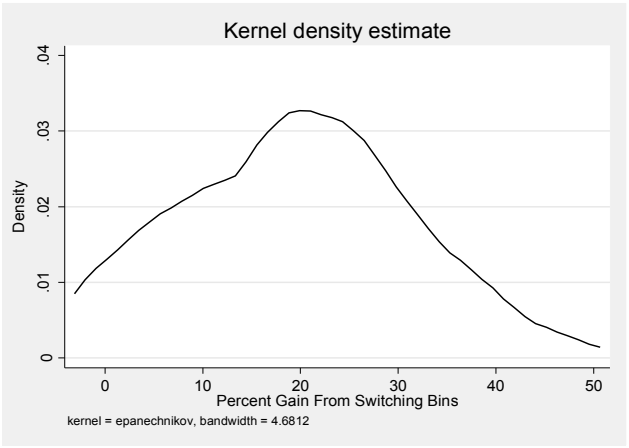
Panel A: Percent Gain from Moving from the Baseline Average to the Recommendation in the Sort Treatment



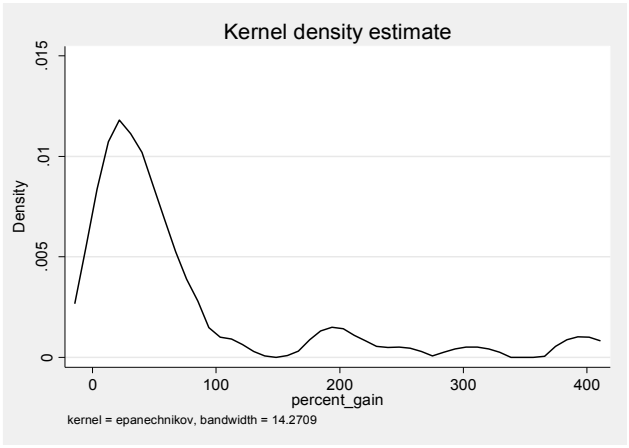


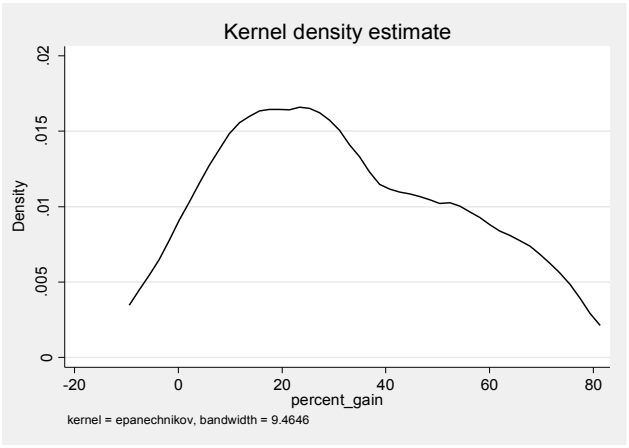
Panel B: Percent Gain from Moving from the Lowest Performing to the Highest Performing Bin in the Sort Treatment





Panel C: Percent Gain from Moving from the Current Pod Size to the Best in the Weight Treatment





Notes: The first column of figures provides the full set of returns, while the second column focuses on the bottom 80 percent of the sample.

Appendix Table 1: Regression Check

	Mean, by Treatment Group			Differences		
	Control	Sort	Weight	Col 1 - Col 2	Col 1 - Col 3	Col 2 - Col 3
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Monthly Per Capita Income)	12.43	12.42	12.67	0.233	-0.016	-0.249
	(1.37)	(0.72)	(0.76)	(0.164)	(0.156)	(0.138)*
HH Head is Literate	0.79	0.84	0.88	0.093	0.050	-0.043
	(0.41)	(0.37)	(0.32)	(0.059)	(0.060)	(0.064)
Number of Assets	8.36	7.95	8.28	-0.079	-0.409	-0.330
	(3.15)	(3.19)	(3.06)	(0.511)	(0.494)	(0.580)
Age of HH Head	43.23	43.95	43.54	0.304	0.718	0.414
	(12.22)	(12.14)	(11.33)	(1.937)	(1.905)	(2.192)
Years Farming	18.00	18.94	18.27	0.275	0.937	0.662
	(6.96)	(7.06)	(7.06)	(1.183)	(1.103)	(1.330)
Parents Farmed Seaweed	0.47	0.55	0.53	0.063	0.081	0.019
	(0.50)	(0.50)	(0.50)	(0.083)	(0.078)	(0.093)
Loans from Someone Sells to	0.31	0.33	0.22	-0.085	0.026	0.111
	(0.46)	(0.48)	(0.42)	(0.088)	(0.085)	(0.097)
Farms Cottoni	0.34	0.33	0.47	0.135	-0.008	-0.144
	(0.47)	(0.47)	(0.50)	(0.082)*	(0.074)	(0.091)
Podsize at Baseline	109.01	102.83	112.78	3.777	-6.174	-9.951
	(28.79)	(22.57)	(38.00)	(5.850)	(3.884)	(5.930)*
Number of Days of Previous Cycle	36.61	36.00	37.08	0.463	-0.612	-1.075
	(8.24)	(7.31)	(6.54)	(1.178)	(1.190)	(1.281)
P-Value from Joint Test				0.6945	0.6077	0.4027

Notes: This table provides a check on the randomization. Columns 1 - 3 provide the mean and standard deviation of each baseline characteristic for the control group, sort group, and weight group, respectively. Columns 4 - 6 give the difference in means (and standard errors) between the noted experimental groups. The p-value for the joint significance of the treatment on all dependent variables is provided in Columns 4 – 6. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table 2: Reasons Why Farmer May Want to Try a New Method

	Mean	Standard Deviation	N
	(1)	(2)	(3)
Would Not Want to Make any Changes	0.04	0.18	482
Own Initiative	0.02	0.13	482
Pest or Price	0.02	0.13	482
Advice from Friend	0.10	0.30	482
NGO or Government Recommendation	0.11	0.31	482
Seeing Results on the Plots of Other Farmers	0.39	0.49	482

Notes: This table provides baseline survey responses on why a farmer may try a new method.

Appendix Table 3: Effect of Treatment, by Recommendation

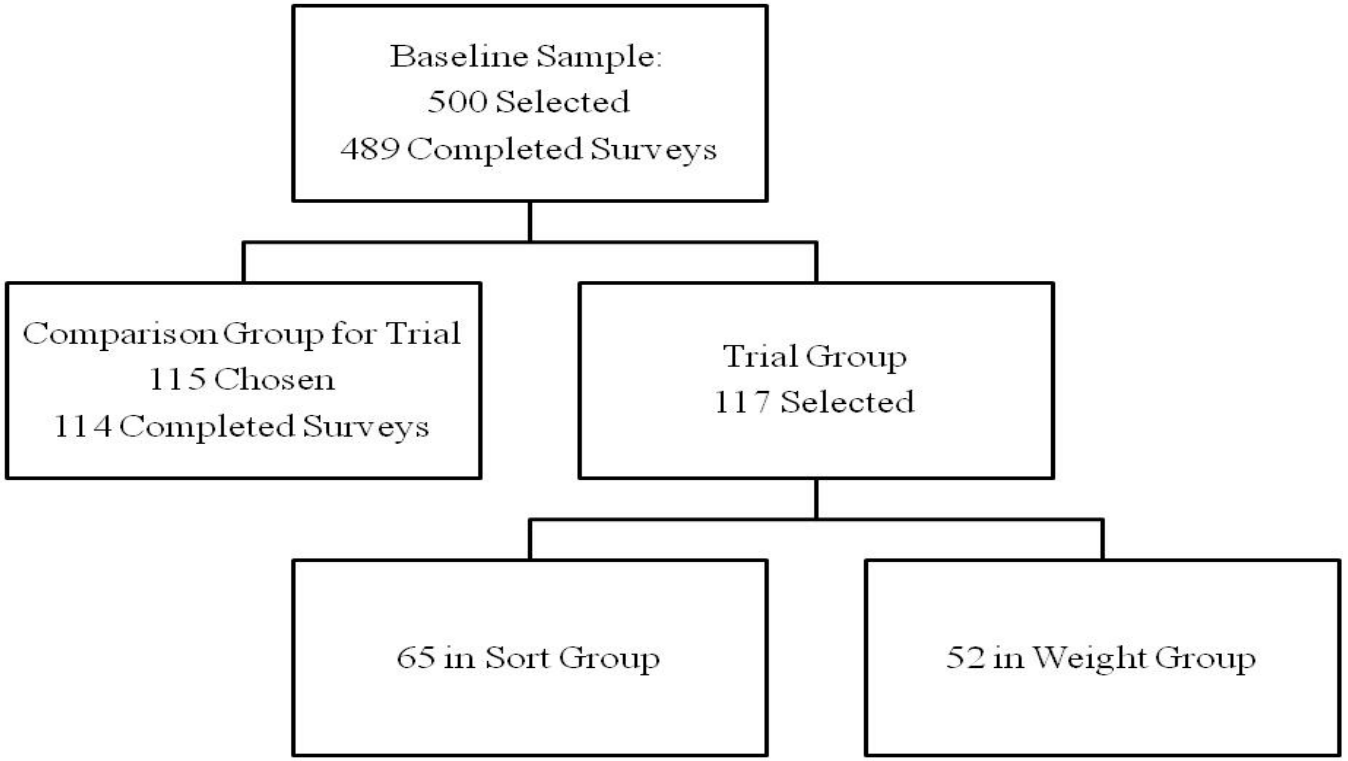
	Pod Size	
	(1)	(2)
Increase	-21.091	-50.976
	(4.807)***	(1.851)***
Increase * Trial Participation * After Trial	27.264	27.060
	(6.223)***	(7.627)***
Increase * Trial Participation * After Summary Data	19.502	19.242
	(5.808)***	(7.149)***
Hamlet Fixed Effects	X	
Farmer Fixed Effects		X
Observations	675	675

Notes: This table provides the coefficient estimates of the effect of the recommendations from the trial on farming methods after the trial (follow-up 1) and after seeing the summary data (follow-up 2), conditional on baseline farming methods. The trial participation dummy indicates that the farmer belongs in either the sort or weight treatment group. Increase is an indicator variable for being told to increase size. All regressions are estimated using OLS and include the main effects as well as the double interactions, and standard errors are clustered at the farmer level. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

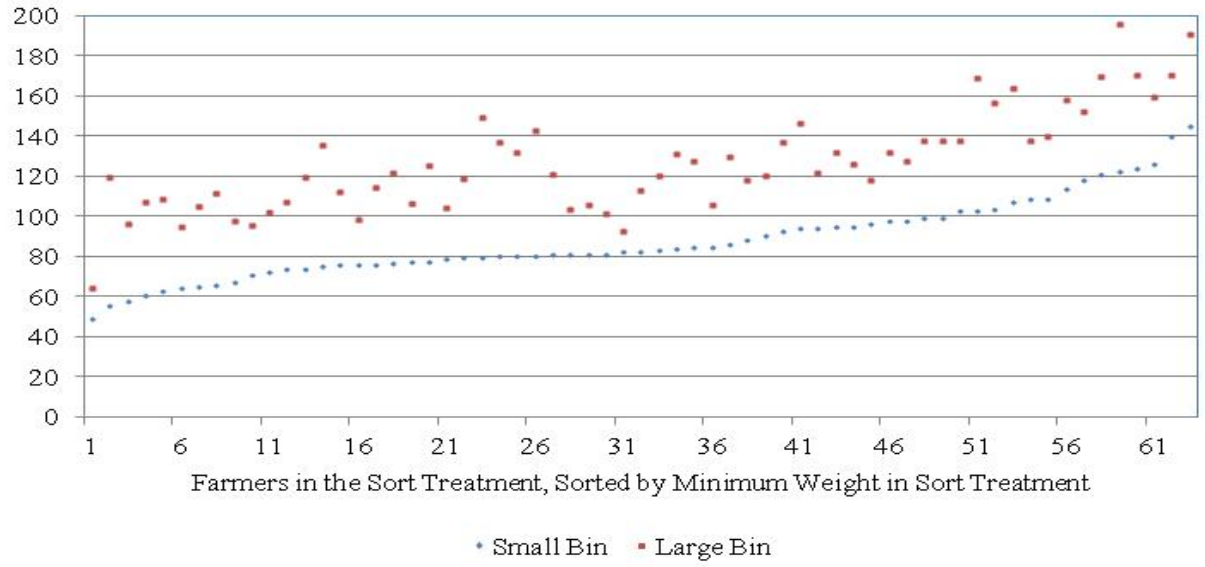
Appendix Figure 1: Experimental Design

	Jun 2007	Jul 2007	Aug 2007	Sept 2007	Oct 2007	Nov 2007	Dec 2007	Jan 2008	Feb 2008	Mar 2008	Apr 2008	May 2008	Jun 2008	Jul 2008	Aug 2008
Baseline															
Trials															
Follow-Up 1															
Follow-Up 2															

Appendix Figure 2: Sample Design



Appendix Figure 3



Notes: This figure documents the within farmer variation in pod sizes at baseline. For each farmer, it shows the average size within the small and large bin in the sort treatment, where farmers are sorted according to the average size in their small bin.

Appendix Figure 4: Examples of Summarized Trial Results

A. Weight Example

Pod Size	Distance	#Pods per line	Initial investment	Return per line
40	15	33	1650	4510
40	20	26	1300	3553
60	15	33	2310	1517
60	20	26	1820	1195
80	15	33	2970	1871
80	20	26	2340	1474
100	15	33	3630	1904
100	20	26	2860	1500
120	15	33	4290	597
120	20	26	3380	470
140	15	33	4950	1574
140	20	26	3900	1240

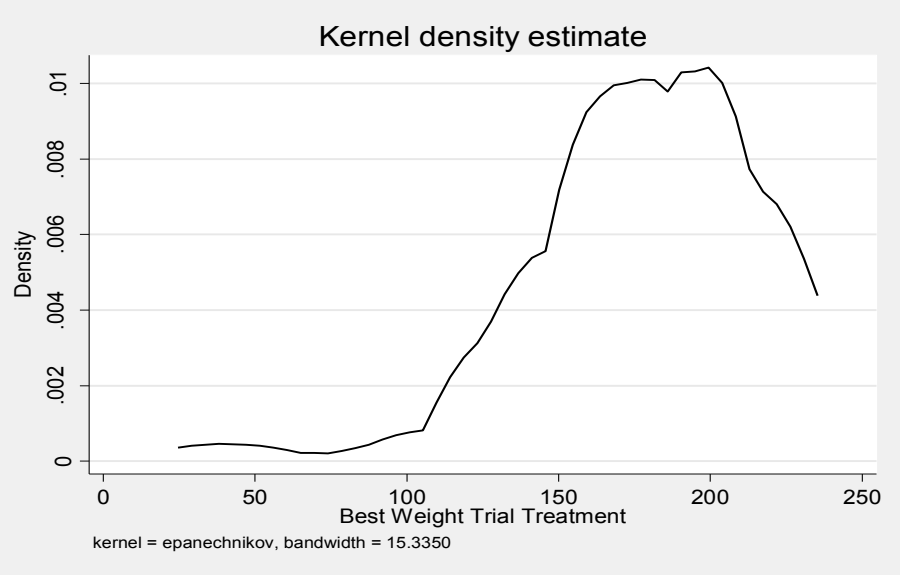
Currently	
Pod Weight: 152.5	**Distance: 15**
Recommendation:	
Pod Weight: 40	**Distance: 15**

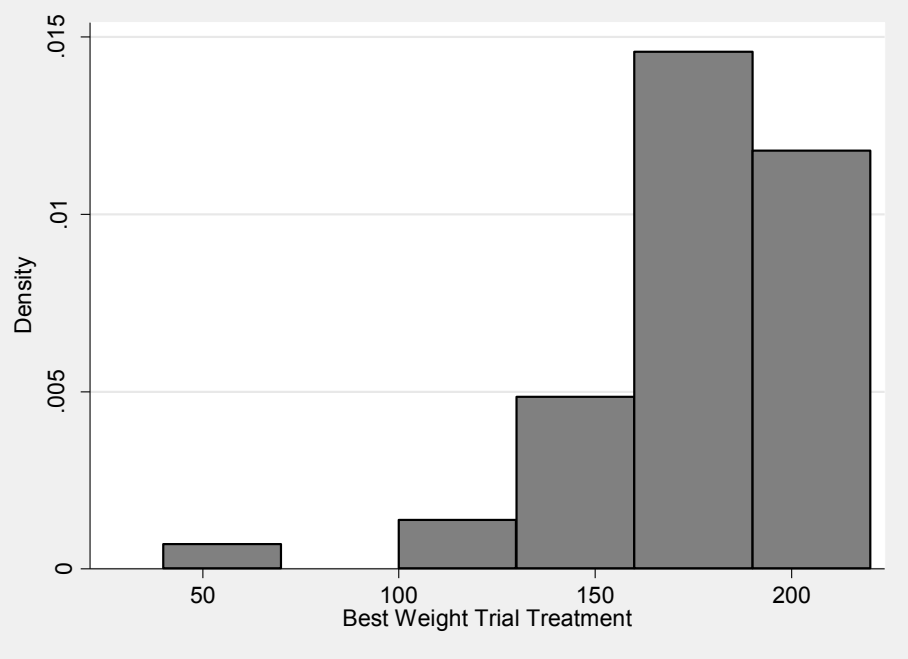
B. Sort Example

Line Type	Distance	Average Pod Weight (g)	Return per pod (g)	Average pods per line
Large	14	129.92	167.65	34.1
Medium		98.34	155.51	
Small		86.18	158.82	

Recommendation	Switch to **large** pod size, with average weight **129.92**g.

Appendix Figure 5: Distribution of Recommended Pod Sizes in the Weight Treatment





References

Anderson, J.R. and G. Feder, "Agricultural Extension: Good Intentions and Hard Realities," *The World Bank Research Observer*, 2004, *19* (1), 41–60.

Beaman, L., J. Magruder, and J. Robinson, "Minding Small Change: Limited Attention Among Small Firms in Kenya," 2012.

Besley, T. and A. Case, "Modeling Technology Adoption in Developing Countries," *American Economic Review*, 1993, *83* (2), 396–402.

____ and ____ , "Diffusion as a Learning Process: Evidence from HYV Cotton," *Working Papers*, 1994.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts, "Does Management Matter: Evidence from India," 2012.

Bordalo, P., N. Gennaioli, and A. Shleifer, "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 2012, *Forthcoming.*

Brandner, L. and M.A. Strauss, "Congruence Versus Profitability in the Diffusion of Hybrid Sorghum," *Rural Sociology*, 1959, *24* (4), 381–383.

Conley, T.G. and C.R. Udry, "Learning About a New Technology: Pineapple in Ghana," *American Economic Review*, 2010, *100* (1), 35–69.

Croppenstedt, A., M. Demeke, and M.M. Meschi, "Technology Adoption in the Presence of Constraints: The Case of Fertilizer Demand in Ethiopia," *Review of Development Economics*, 2003, *7* (1), 58–70.

Duflo, E., M. Kremer, and J. Robinson, "Why Don't Farmers Use Fertilizer? Experimental Evidence from Kenya," *Massachusetts Institute of Technology and Harvard University, working paper*, 2008.

Evenson, R.E., "Economic Impacts of Agricultural Research and Extension," *Handbook of Agricultural Economics*, 2001, *1*, 573–628.

Fliegel, F.C. and J.E. Kivlin, "Attributes of Innovations as Factors in Diffusion," *American Journal of Sociology*, 1966, pp. 235–248.

Foster, A.D. and M.R. Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 1995, pp. 1176–1209.

___ and ___, "Microeconomics of Technology Adoption," *Annual Reviews of Economics*, 2010, *2* (1), 395–424.

Gabaix, X., "A Sparsity-Based Model of Bounded Rationality," 2011.

Gautam, M., *Agricultural Extension: The Kenya Experience: An Impact Evaluation*, World Bank Publications, 2000.

Gawande, A., "On Washing Hands," *New England Journal of Medicine*, 2004, *350* (13), 1283–1286.

Gerhart, J.D., "The Diffusion of Hybrid Maize in Western Kenya," 1975.

Gittins, J.C., "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1979, pp. 148–177.

Griliches, Z., "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 1957, pp. 501–522.

Jovanovic, B. and Y. Nyarko, "Learning by Doing and the Choice of Technology," *Econometrica*, 1996, *64* (6), 1299–1310.

Koszegi, B. and A. Szeidl, "A Model of Focusing in Economic Choice," 2011.

Nelson, R.R. and E.S. Phelps, "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, 1966, *56* (1/2), 69–75.

Niehaus, P., "Filtered Social Learning," *Journal of Political Economy*, 2011, *119* (4), 686–720.

Nuland, S.B., *The Doctors' Plague: Germs, Childbed Fever, and the Strange Story of Ignac Semmelweis*, WW Norton & Company, 2004.

Rogers, E.M. and F.F. Shoemaker, "Communication of Innovations; A Cross-Cultural Approach.," 1971.

Rosenzweig, M.R., "Why are There Returns to Schooling?," *American Economic Review*, 1995, *85* (2), 153–158.

Schultz, T.W., *The Economic Value of Education*, Vol. 63, Columbia University Press New York, 1963.

____ , "The Value of the Ability to Deal with Disequilibria," *Journal of Economic Literature*, 1975, *13* (3), 827–846.

Schwartzstein, J., "Selective Attention and Learning," *Unpublished Manuscript, Dartmouth College*, 2012.

Sims, C.A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Suri, T., "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 2011, *79* (1), 159–209.

Tornatzky, L.G. and K.J. Klein, "Innovation Characteristics and Innovation Adoption-Implementation: A Meta-Analysis of Findings," *IEEE Transactions on Engineering Management*, 1982, *29* (1), 28–45.

Zwane, A.P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D.S. Karlan, R. Hornbeck, X. Giné et al., "Being Surveyed can Change Later Behavior and Related Parameter Estimates," *Proceedings of the National Academy of Sciences*, 2011, *108* (5), 1821–1826.

A Proofs

Before turning to the proofs, we introduce some notation and a useful lemma.

Across the regimes studied in Section 2, in the second period the farmer will compare

$$\mu^{+*} \equiv \max_{s' \in \mathcal{S}} E[f(s')|\hat{h}] = \max_{s' \in \mathcal{S}} \pi^{+} E[f(s')|\hat{h}, P] + (1 - \pi^{+}) E[f(s')|\hat{h}, NP] \tag{12}$$

with

$$\bar{\mu}^{+} \equiv \frac{1}{S} \sum_{s'} E[f(s')|\hat{h}] \tag{13}$$

and, for every piece of land, will choose to apply the input level s^* that achieves μ^{+*} if $\mu^{+*} - \bar{\mu}^{+} > m$ and will not measure otherwise. Label the event that $\mu^{+*} - \bar{\mu}^{+} > m$ by G and the event that $\mu^{+*} - \bar{\mu}^{+} \leq m$ by NG. Further, let

$$\mu_M^{+*} \equiv E[f(s^*)|\hat{h}, M] \tag{14}$$

and

$$\bar{\mu}_M^{+} \equiv \frac{1}{S} \sum_{s'} E[f(s')|\hat{h}, M] \tag{15}$$

for $M = P, NP$. Note that $\mu_{NP}^{+*} = \bar{\mu}_{NP}^{+}$ given any history and that $\mu_P^{+*} - \bar{\mu}_P^{+} > m/\pi^{+}$ conditional on event G.

Lemma A.1. *Independent of the regime studied in Section 2:*

1. *If the farmer does not keep track of the input level applied to any piece of land in the first period, then $\mu^{+*} = \bar{\mu}^{+}$ along any history.*

2. *If the farmer keeps track of the input level in the first period, then the probability of event G tends towards 1 as m tends towards 0 under model P.*

Proof. **Part 1.** We have

$$\begin{aligned}\mu^{+*} &= \max_{s'} \pi^{+} E[f(s')|\hat{h}, P] + (1-\pi^{+}) E[f(s')|\hat{h}, NP] \\ &= \max_{s'} \pi^{+} E[\theta(s')|\hat{h}, P] + (1-\pi^{+}) E[\theta|\hat{h}, NP],\end{aligned} \tag{16}$$

and

$$\begin{aligned}\bar{\mu}^{+} &= \pi^{+} \left(\frac{1}{S} \sum_{s'} E[f(s')|\hat{h}, P] \right) + (1-\pi^{+}) \left(\frac{1}{S} \sum_{s'} E[f(s')|\hat{h}, NP] \right) \\ &= \pi^{+} \left(\frac{1}{S} \sum_{s'} E[\theta(s')|\hat{h}, P] \right) + (1-\pi^{+}) \left(\frac{1}{S} \sum_{s'} E[\theta|\hat{h}, NP] \right).\end{aligned} \tag{17}$$

Comparing (16) with (17), we see that it suffices to show that

$$E[\theta(s')|\hat{h}, P] = E[\theta(s'')|\hat{h}, P] \tag{18}$$

for all $s', s'' \in \mathcal{S}$ and $\hat{h}$ consistent with the farmer not keeping track of the input in the first period. But this equality follows from the fact that the farmer's marginal prior distributions over $\theta(s')$ and $\theta(s'')$ are the same under model P, as are the likelihoods $\Pr(\hat{h}|\theta(s') = \tilde{\theta}, P)$ and $\Pr(\hat{h}|\theta(s'') = \tilde{\theta}, P)$ for all $s', s'', \tilde{\theta}$, and $\hat{h} = (y_{p1}, \varnothing)_{p=1}^{N}$.

Part 2.

We will consider the learning by doing regime. The proof of this result for the other regimes is similar but slightly less involved because, in those regimes, $n(s)$ is known rather than stochastic.

Define $\mu_P^s = E[f(s)|\hat{h}, P]$ for all $s \in \mathcal{S}$, so $(\mu_P^1, \mu_P^2, \ldots, \mu_P^S)$ is a random variable. Re-order these μ_P^j in terms of decreasing values, where $\mu_P^{(i)}$ denotes the i-th highest value, so $\mu_P^{(1)} \geq \mu_P^{(2)} \geq \ldots \geq \mu_P^{(S)}$. Clearly, $(\mu_P^{(1)}, \mu_P^{(2)}, \ldots, \mu_P^{(S)})$ is also a random variable.

By definition, event NG is the event that $\mu^{+*} - \bar{\mu}^{+} \leq m$ which is equivalent to the event that

$$\mu_P^{+*} - \bar{\mu}_P^{+} \leq m/\pi^{+}. \tag{19}$$

It turns out that (19) holds only if

$$|\mu_P^s - \mu_P^{s'}| \leq \frac{Sm}{\pi^+} \text{ for all } s, s'. \tag{20}$$

To see this, suppose that there exists some s, s' such that $|\mu_P^s - \mu_P^{s'}| > Sm/\pi^+$. Then, for some $j > 0, k > 0$, we must have $\mu_P^{(j)} - \mu_P^{(j+k)} > Sm/\pi^+$. Now, expanding $\mu_P^{+*} - \bar{\mu}_P^+$, we have

$$\begin{aligned}
\mu_P^{+*} - \bar{\mu}_P^+ &= \mu_P^{(1)} - \frac{1}{S}\left(\mu_P^{(1)} + \mu_P^{(2)} + \ldots + \mu_P^{(j)} + \ldots + \mu_P^{(j+k)} + \ldots + \mu_P^{(S)}\right) \\
&\geq \mu_P^{(1)} - \frac{1}{S}\left(\underbrace{\mu_P^{(1)} + \ldots + \mu_P^{(1)}}_{(j-1)\text{-times}} + \underbrace{\mu_P^{(j)} + \ldots + \mu_P^{(j)}}_{k\text{-times}} + \underbrace{\mu_P^{(j+k)} + \ldots + \mu_P^{(j+k)}}_{S-(j-1)-k\text{-times}}\right) \\
&= \frac{S-(j-1)}{S}\mu_P^{(1)} - \frac{k}{S}\mu_P^{(j)} - \frac{S-(j-1)-k}{S}\mu_P^{(j+k)} \\
&\geq \frac{S-(j-1)-k}{S}(\mu_P^{(j)} - \mu_P^{(j+k)}) \\
&> \frac{S+1-(j+k)}{S}\left(\frac{Sm}{\pi^+}\right) \\
&\geq \frac{1}{S}\left(\frac{Sm}{\pi^+}\right) \\
&= \frac{m}{\pi^+}.
\end{aligned}$$

Having established that (19) holds only if (20), we will now show that

$$\lim_{m\to 0} \Pr\left(|\mu_P^s - \mu_P^{s'}| \leq \frac{Sm}{\pi^+} \text{ for all } s, s'|P\right) = \lim_{m'\to 0} \Pr\left(\pi^+|\mu_P^s - \mu_P^{s'}| \leq m' \text{ for all } s, s'|P\right) = 0. \tag{21}$$

Note that

$$\Pr\left(\pi^+|\mu_P^s - \mu_P^{s'}| \leq m' \text{ for all } s, s'|P\right) \leq \Pr\left(\pi^+|\mu_P^s - \mu_P^{s'}| \leq m' \text{ for some } s, s' \text{ with } n(s) > 0|P\right), \tag{22}$$

so it is enough to establish that the right hand side of this inequality tends towards zero as $m' \to 0$.

By Bayes' rule, the right hand side of (22) can be expanded as follows (where we let $Q(\cdot)$ denote $\Pr(\cdot|n(s) > 0, P)$):

$$Q\left(\pi^+|\mu_P^s - \mu_P^{s'}| \leq m'|\pi^+ \geq \underline{\pi}\right)\Pr(\pi^+ \geq \underline{\pi}|P) + Q\left(\pi^+|\mu_P^s - \mu_P^{s'}| \leq m'|\pi^+ < \underline{\pi}\right)\Pr(\pi^+ < \underline{\pi}|P)$$

for any $\underline{\pi} \in (0,1)$, which is less than or equal to $Q\left(\underline{\pi}|\mu_P^s - \mu_P^{s'}| \leq m'|\pi^+ \geq \underline{\pi}\right)\Pr(\pi^+ \geq \underline{\pi}|P) + \Pr(\pi^+ < \underline{\pi}|P)$, which in turn is less than or equal to

$$Q\left(\underline{\pi}|\mu_P^s - \mu_P^{s'}| \leq m'|\pi^+ \geq \underline{\pi}\right)(1 - p(\underline{\pi})) + p(\underline{\pi}) \tag{23}$$

where $p(\underline{\pi}) = \frac{1-\pi}{\pi} \cdot \frac{\underline{\pi}}{1-\underline{\pi}}$. The last inequality follows from the fact that the odds ratio $(1-\pi_t)/\pi_t$ is a supermartingale under model P, together with Markov's inequality.[32]

Conditional on $n(s) > 0$, $|\mu_P^s - \mu_P^{s'}|$ is a continuous random variable under model P on $[0, \infty)$ (μ_P^s is a normally distributed random variable and $\mu_P^{s'}$ is either an independent normal random variable or a constant), so

$$\lim_{\tilde{m} \to 0} \Pr\left(|\mu_P^s - \mu_P^{s'}| \leq \tilde{m}|n(s) > 0, P\right) = 0,$$

which implies that (23) tends to $p(\underline{\pi})$ as $m' \to 0$ for any $\underline{\pi}$. Since it is additionally the case that $p(\underline{\pi}) \to 0$ as $\underline{\pi} \to 0$, we have established (21), and hence that the probability of event G tends towards 1 as m tends towards 0.

■

Proof of Proposition 1. Label the farmer's first period choice of action by a, which, for every piece of land, specifies whether the farmer measures and/or attends to the input level applied to that piece of land. The farmer's first period choice of action, together with his prior and second

[32] Markov's inequality says that for any random variable X and $a > 0$, $\Pr(|X| \geq a) \leq E[|X|]/a$. Given that $(1-\pi_t)/\pi_t$ is a supermartingale under model P, we have that $\Pr\left(\frac{1-\pi^+}{\pi^+} \geq a|P\right) \leq \frac{1-\pi}{\pi a}$, which in turn implies that $\Pr\left(\pi^+ \leq \frac{1}{1+a}|P\right) \leq \frac{1-\pi}{\pi a}$, or, changing variables, $\Pr\left(\pi^+ \leq \underline{\pi}|P\right) \leq \frac{1-\pi}{\pi} \cdot \frac{\underline{\pi}}{1-\underline{\pi}}$.

period strategy (as detailed at the beginning of the Appendix) induces a probability distribution over outcomes. Label this distribution by $\Pr^a$.

Given action a, the farmer's first period expected flow utility deriving from piece of land p is

$$\mu - \delta^a_{pm} m - \delta^a_{pe} e, \tag{24}$$

where δ^a_{pm} is an indicator for whether the input is measured, and δ^a_{pe} is an indicator for whether the farmer keeps track of the input (applied to piece of land p under action a).

Conditional on the farmer's first period action, his second period expected (per-piece of land) flow utility under model NP is

$$\begin{aligned} u^a_{NP} &= \Pr^a(G|NP)(E^a[\mu^{+*}_{NP}|G, NP] - m) + \Pr^a(NG|NP)E^a[\bar{\mu}^+_{NP}|NG, NP] \\ &= E^a[\bar{\mu}^+_{NP}|NP] - m \Pr^a(G|NP) \text{ (since } \mu^{+*}_{NP} = \bar{\mu}^+_{NP}) \\ &= \mu - m \Pr^a(G|NP) \text{ (by the law of iterated expectations).} \end{aligned} \tag{25}$$

The farmer's second period expected (per-piece of land) flow utility under model P is

$$u^a_P = \Pr^a(G|P)(E^a[\mu^{+*}_P|G, P] - m) + \Pr^a(NG|P)E^a[\bar{\mu}^+_P|NG, P]. \tag{26}$$

It is easy to see that

$$\mu - m \leq u^a_{NP} \leq \mu \tag{27}$$

and

$$\mu \leq u^a_P \leq E^a[\mu^{+*}_P|P]. \tag{28}$$

A less obvious, but important inequality is that, conditional on the agent keeping track of the input

level in the first period, there exists an $\varepsilon > 0$ such that

$$u_P^a > \mu + \varepsilon \tag{29}$$

for sufficiently small m. We will take (29) as given for now, and will establish this inequality at the end of the proof.

The farmer's expected utility given action a is

$$U^a = N\mu - \sum_{p=1}^{N}(\delta_{pm}^a m + \delta_{pe}^a e) + N[(1-\pi)u_{NP}^a + \pi u_P^a]. \tag{30}$$

When the farmer does not measure or keep track of the input level applied to any piece of land ($a = NE$), Lemma A.1 gives us that $\Pr^{NE}(G) = 0$, which implies that $u_P^{NE} = u_{NP}^{NE} = \mu$. Thus, the farmer's expected utility when he does not keep track of the input level equals

$$U^{NE} = 2N\mu. \tag{31}$$

If the farmer takes some other action (i.e., one in which he does keep track of the input level) his expected utility is bounded above by

$$N\{\mu + (1-\pi)\mu + \pi \max_a E^a[\mu_P^{+*}|P]\} - e \tag{32}$$

by (27), (28), and the fact that $e \leq m$. It is easy to see that (32) is less than U^{NE} for sufficiently small π or sufficiently large e. Thus, the farmer chooses not to measure or keep track of the input level for sufficiently small π or sufficiently large e, m. When the farmer does not measure or keep track of the input level, Lemma A.1 tells us that his posterior expectation of $f(s)$ will not vary in s and he chooses not to measure in the second period.

We now want to show that the farmer keeps track of the input level applied to some piece of land in the first period so long as π is sufficiently large and/or e and m are sufficiently low. The

utility the farmer achieves through the best strategy in which he keeps track of some input level in the first period is bounded below by the utility from the strategy that includes not measuring, yet attending to the input level applied to every piece of land in the first period. Label the expected utility from this strategy by U^E. We have that

$$\begin{aligned} U^E &= N[\mu - e + (1-\pi)u^E_{NP} + \pi u^E_P] \\ &\geq N[\mu - e + (1-\pi)(\mu - m) + \pi u^E_P] \text{ (by inequality (27))} \\ &> N[2\mu - e - (1-\pi)m + \pi\varepsilon] \text{ for sufficiently low } m\text{(by inequality (29)).} \end{aligned} \tag{33}$$

The right hand side of the above inequality is increasing in π and decreasing in m and e. Further, the right hand side exceeds U^{NE} when π is sufficiently large (fixing sufficiently low e, m) or e and m are sufficiently low (fixing π), thus establishing that the farmer keeps track of the input level applied to a piece of land in the first period so long as π is sufficiently large (fixing sufficiently low e, m) or e and m are sufficiently low (fixing π). When he keeps track of the input level applied to a given piece of land, his posterior expectation almost surely varies in s when $\sigma^2_\varepsilon > 0$ because yield is normally distributed given s' and $f_0(s')$, and is not perfectly correlated across pieces of land. It is obvious that, in this case, the farmer either measures the input level with the highest associated posterior distribution of $f(s)$ in the second period or does not measure.

It only remains to show that (29) holds. Let G^δ be the event that $\mu^{+*} - \bar{\mu}^+ > m + \delta$ and $G^{\neg\delta}$ be the event that $m < \mu^{+*} - \bar{\mu}^+ \leq m + \delta$. We can write

$$\begin{aligned} u^a_P &= \Pr^a(G^\delta|P)(E^a[\mu^{+*}_P|G^\delta, P] - m) + \Pr^a(G^{\neg\delta}|P)(E^a[\mu^{+*}_P|G^{\neg\delta}, P] - m) \\ &\quad + \Pr^a(NG|P)E^a[\bar{\mu}^+_P|NG, P] \\ &\geq \mu + \Pr^a(G^\delta|P)\delta, \end{aligned} \tag{34}$$

with strict inequality whenever $\Pr^a(G|P) > 0$. The inequality follows from the fact that $\mu^{+*}_P - \bar{\mu}^+_P > m + \delta$ conditional on event G^δ, together with the law of iterated expectations.

Lemma A.1 implies that, when the farmer keeps track of the input level in the first period, there

exists some $p, \delta, \bar{m}$ such that $\Pr^a(G^\delta|P) \geq p$ for all $m \leq \bar{m}$, so (34) implies that $u_P^a > \mu + p\delta$ for all $m \leq \bar{m}$. Setting $\varepsilon = p\delta$ completes the proof.

■

Proof of Proposition 2. Using arguments and notation analogous to those used in the proof of Proposition 1, the farmer's expected utility if he does not attend to the input level applied to any piece of land in the first period equals

$$U^{NA} = 2N\mu, \tag{35}$$

while his expected utility if he attends to some input level is bounded above by

$$N\{\mu + (1-\pi)\mu + \pi \max_a E^a[\mu_P^{+*}|P]\} - e. \tag{36}$$

Since (36) is less than U^{NA} for sufficiently small π or sufficiently large e, the farmer does not attend to the input level for sufficiently small π or large e, m in the learning by observing regime. Thus, a farmer who would not have attended to the input level absent exogenous experimentation also may not in this regime. Finally, such a farmer's posterior expectation of $f(s)$ does not vary in s by Lemma A.1.

■

Before proving Proposition 3, we state and prove an intermediate Lemma.

Lemma A.2. *Assume that $\sigma_\varepsilon^2 = 0$. Given any history h, the farmer will make the same decision in the second period whether he knows h, or just $((y_{p1}, \varnothing)_{p=1}^N, \bar{y}_{s^*}, s^*) \equiv \tilde{h}$.*

Proof. Recall that, in the second period, the farmer will choose a level s^* that achieves μ^{+*} if $\mu^{+*} - \bar{\mu}^+ > m$ and will not measure otherwise. It will be helpful to write μ^{+*} and $\bar{\mu}^+$ as functions

of information known by the farmer, so

$$\mu^{+*}(h) = \max_{s' \in \mathcal{S}} E[f(s')|h],$$
$$\mu^{+*}(\tilde{h}) = \max_{s' \in \mathcal{S}} E[f(s')|\tilde{h}],$$

etc. It suffices to show that $\mu^{+*}(h) = \mu^{+*}(\tilde{h})$ and $\bar{\mu}^{+}(h) = \bar{\mu}^{+}(\tilde{h})$ for all h.

We have

$$\begin{aligned}\mu^{+*}(h) &= \max_{s' \in \mathcal{S}} \pi^{+}(h) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\sigma_\varepsilon^2 + n\sigma^2}\bar{y}_{s'} \right) + (1 - \pi^{+}(h)) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + N\sigma^2}\mu + \frac{N\sigma^2}{\sigma_\varepsilon^2 + N\sigma^2}\bar{y} \right) \\ &= \pi^{+}(h) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\sigma_\varepsilon^2 + n\sigma^2}\bar{y}_{s^*} \right) + (1 - \pi^{+}(h)) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + N\sigma^2}\mu + \frac{N\sigma^2}{\sigma_\varepsilon^2 + N\sigma^2}\bar{y} \right).\end{aligned} \tag{37}$$

From (37), we see that, to calculate $\mu^{+*}(h)$, the farmer only needs to know $(s^*, \bar{y}_{s^*}, \bar{y}, \pi^{+}(h))$, which he can calculate from $\tilde{h}$ so long as $\sigma_\varepsilon^2 = 0$. The only part of this statement that requires explanation is why he can calculate $\pi^{+}(h)$ from $\tilde{h}$. This follows from the facts that, given $\sigma_\varepsilon^2 = 0$, we have (i)

$$\pi^{+}(h) = \begin{cases} 1 & \text{if there exists } s, s' \text{such that } \bar{y}_s \neq \bar{y}_{s'} \\ 0 & \text{otherwise} \end{cases}$$

and (ii) there exists s, s' such that $\bar{y}_s \neq \bar{y}_{s'}$ if and only if $\bar{y}_{s^*} \neq \bar{y}$. Thus, $\mu^{+*}(h) = \mu^{+*}(\tilde{h})$.

We also have

$$\begin{aligned}\bar{\mu}^{+}(h) &= \frac{1}{S}\sum_{s'} \left\{ \pi^{+} \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\sigma_\varepsilon^2 + n\sigma^2}\bar{y}_{s'} \right) + (1 - \pi^{+}) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + N\sigma^2}\mu + \frac{N\sigma^2}{\sigma_\varepsilon^2 + N\sigma^2}\bar{y} \right) \right\} \\ &= \pi^{+}(\tilde{h}) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\sigma_\varepsilon^2 + n\sigma^2}\bar{y} \right) + (1 - \pi^{+}(\tilde{h})) \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + N\sigma^2}\mu + \frac{N\sigma^2}{\sigma_\varepsilon^2 + N\sigma^2}\bar{y} \right).\end{aligned} \tag{38}$$

From (38) we see that, to calculate $\bar{\mu}^{+}(h)$, the farmer only needs to know $\bar{y}$ and $\bar{y}_{s^*}$ when $\sigma_\varepsilon^2 = 0$, which can be calculated from $\tilde{h}$. Thus, $\bar{\mu}^{+}(h) = \bar{\mu}^{+}(\tilde{h})$. ■

Remark A.1. From the proof of Lemma A.2, we can see that the "sufficient statistics" statement would continue to hold if, instead of $\sigma_\varepsilon^2 = 0$, the farmer did not update model weights in response to new information (just beliefs given a model).

Proof of Proposition 3. The farmer essentially has four options in the first period:

1. Keep track of the input level applied to some pieces of land.
2. Keep track of the input level applied to some pieces of land, and attend to recommendation.
3. Don't keep track of the input level, but attend to recommendation.
4. Don't keep track of the input level or attend to recommendation.

By Lemma A.2, options 1 and 2 are dominated by 3, so the farmer compares 3, attending to the recommendation ($a = R$), with 4, not attending to the recommendation ($a = NR$).

The farmer's expected utility if he does not attend to the recommendation is $U^{NR} = 2N\mu$ by now familiar arguments.

If the farmer attends to the recommendation, then the probability distribution over his second period action is the same as if he kept track of the input level applied to *every* piece of land by Lemma A.2. However, the cost of attending to the recommendation is only e – the same as the cost of attending to the input level applied to a *single* piece of land. Thus, the farmer's expected utility if he attends to the recommendation, U^R, is greater than his expected utility under the best strategy where he attends to the input level applied to at least one piece of land under the learning by observing regime, U^A.

Proposition 3 is then an immediate implication of the fact that $U^R \geq U^A$. ■

Proof of Proposition 4. Given the symmetric strategy assumption (Assumption 5.1), the farmer has three options in the first period:

1. Do not measure or attend to the input level applied to any piece of land.
2. Do not measure, but attend to the input level applied to every piece of land.

3. Measure a particular input s on every piece of land.

Together with the fact that the farmer's prior is symmetric across input levels and $e < m$, Assumption 5.2 guarantees that option 2 dominates option 3. Under these assumptions, the farmer is guaranteed to learn the optimal input level under option 2 but not option 3, and option 2 is associated with a lower cost than option 3. As a result, the farmer does not measure the input level in the first period.

If the farmer does not measure but attends to the input level in the first period, then he learns the optimal level (by Assumption 5.2) and continues to farm in the second period, using input level s_0.

If the farmer does not attend to the input level in the first period, then $\mu^{+*} = \bar{\mu}^{+} = \frac{1}{S}\sum_{s'} f_0(s')$ (by Assumption 5.2). If the technology is filter congruent, then $\bar{\mu}^{+} > 0$ and the farmer continues to farm in the second period; if the technology is filter incongruent, then $\bar{\mu}^{+} < 0$ and the farmer does not continue to farm.

■

Proof of Proposition 5. If the farmer attends to the input in period 0, then, by Assumption 5.2, he learns for sure that he is guaranteed a (per-piece-of-land) profit of $f_0(s_0) - m > 0$ by farming and using input s_0. Since this profit exceeds the zero profit he would earn by not farming, he will continue to farm forever.

If the farmer does not attend to the input in period 0, then he will not measure the input if he farms in period 1. If the technology is filter congruent, then he learns that he is guaranteed an income of at least $1/S\sum_{s'} f_0(s') > 0$ by farming, and will continue to profitably farm in period 2. If the technology is filter incongruent and he farms but does not attend in period 1, then the farmer's expected profit is negative if he continues to farm in period 2 (since $\sum_{s'} f_0(s') < 0$), so he stops farming.

■