Selective Trials, Information Production and Technology Diffusion

PRELIMINARY AND INCOMPLETE*

Sylvain Chassang — Princeton University

Gerard Padro i Miquel — London School of Economics

Erik Snowberg — Caltech

August 20, 2009

Abstract

This paper explores how to run randomized experiments when outcomes depend significantly on unobserved effort decisions taken by agents. Our first set of results show that making selection explicit rather than implicit can generate greater information. This information comes at the cost of oversampling agents with high valuation for the treatment and undersampling agents with low valuation. These sampling costs disappear if the sample size is large, or agents are very responsive to incentives. Our second set of results study the effect of information on future adoption. Because effort responds to beliefs, which in turn respond to information, informative experiments can potentially encourage the adoption of useful technologies. We show that while information can increase the expected take-up of a high-return technology under fairly broad conditions, this is not always the case. Whether more information increases usage depends on the quality of the information, on the shape of the mapping between agents' beliefs and usage, and on how agents interpret data.

KEYWORDS: program evaluation, randomized controlled trials, mechanism design, selection, heterogeneous beliefs.
JEL: C81, C93, D82, O12.

*Comments and feedback are highly appreciated. Please do not quote or circulate without contacting the authors. The paper will be updated extensively in the near future.

1 Introduction

This paper explores technology evaluation and technology diffusion when outcomes depend significantly on unobserved effort decisions taken by agents.[1] Even in a standard randomized controlled trial (RCT), where the principal randomly and independently assigns experimental subjects to either receive the technology or to be in the control group, unobserved effort limits the information that can be obtained.[2] For example, if a technology's measured returns are low, it is difficult to disentangle whether this is because the true returns are low or because most agents put no effort into using the technology. Furthermore, given that effort responds to beliefs, and beliefs respond to information, the returns measured by an initial RCT will generally differ from the returns that better informed agents would obtain.

We take a mechanism design approach to examine the following questions:

Information Production: How can trials be designed to extract more information on the returns to a technology?

Technology Diffusion: How does the information generated by trials affect the take up of high return technologies?

The model considers a stylized two-stage experimental setting. To maximize the usefulness of our theory for practice, we put particular emphasis on generality and allow the principal and agents to hold general, and in particular heterogenous, beliefs.[3] In the first stage, where we study *information production*, the principal designs a mechanism that assigns agents to the treatment or control group with a probability that depends on the transfers each agent

[1]Throughout the paper we call experimental subjects agents, and call the experimenter the principal. Following usual conventions, we refer to the principal as she and refer to an agent as he.

[2]See Duflo, Glennerster, and Kremer (2008) for a more detailed description of RCTs. Our contribution here is wholly theoretical, but along the lines of Besley and Case (1993) and Banerjee, Bardhan, Basu, Kanbur, and Mookherjee (2005) we approach theory as an important input for empirical work.

[3]This is particularly important when we study technology diffusion, as heterogenous beliefs will lead agents to interpret data differently. For studies of heterogeneous priors in various economic environments see Piketty (1995), Yildiz (2003), Van den Steen (2004, 2005), Gentzkow and Shapiro (2006), Sandroni and Squintani (2007), Acemoglu, Chernozhukov, and Yildiz (2007) and Sethi and Yildiz (2009).

is willing to make.[4] Each agent in the treatment group then makes an unobserved effort decision, and an outcome is drawn from a distribution that depends on the technology, the agent's effort, and the agent's type. At the end of the first stage, the data generated by the experiment becomes public, i.e. the message sent, the transfer made, and the outcome obtained by each agent are observed by the principal and all other agents. In the second stage, where we study *technology diffusion*, agents use this data to update their beliefs about the technology's returns. All agents are then given free access to the technology and make a private usage decision.

Our first set of results speak to information production, and show that what we call *selective trials* are more informative than randomized controlled trials (RCTs). Selective trials are simple dominance-solvable mechanisms in which agents can signal their value for the technology by selecting a higher likelihood of treatment at some cost.[5] These selective trials are best seen as a straightforward extension of RCTs in which likelihood of treatment can vary with agents. We also show that the agents' value for the treatment is the only additional information that can be elicited in dominant strategies. In that sense selective trials are maximally informative.

Selective trials build on techniques already used in the field (see e.g. Ashraf, Berry, and Shapiro, 2008; Karlan and Zinman, 2009), and may be useful in mitigating some of the existing concerns over RCTs (Deaton, 2009). RCTs recover the local average treatment effect (LATE) of the technology (Imbens and Angrist, 1994). However, subjects assigned to the treatment group may refuse to be treated, and those assigned to the control group may

[4]Transfers can be thought of as monetary payments or, more generally, anything that requires the agent to exert *observable* costly effort. For example, agents could choose between lines of different lengths to place themselves into the treatment group with different probabilities.

[5]The idea that a higher price will select individuals with a higher value has been in the economics literature for some time as it is closely related to classic selection models, see Roy (1951) and Oster (1995). Recently Ashraf, Berry, and Shapiro (2008) have found support for this hypothesis in a study evaluating a water treatment product in Zambia. In contrast, Cohen and Dupas (2009) and Dupas (2009b) find that price has no effect on usage in a study evaluating mosquito nets for malaria prevention in Kenya.

adopt the technology.[6] Further, when usage is imperfectly observable the principal may not even be aware of these subversions of the experimental protocol.[7]

Selective trials mitigate these problems by allowing agents to explicitly select themselves into the treatment and control group based on their beliefs and their valuation of the technology.[8] Such selective trials allow the principal to estimate the returns to the technology for agents that value it differently. This corresponds to identifying the marginal treatment effects analyzed by Heckman and Vytlacil (2005). Moreover, by offering the agents explicit ways to select themselves in or out of the treatment, selective trials provide useful information about the agents' unobserved effort decisions. The basic insight underlying our approach is that when unobservable effort cannot be controlled, explicit selection is preferable to implicit selection.

It is important to note that selective trials inherit many of the robustness properties that make RCTs attractive. In particular, the mechanisms we discuss identify the agents' values for the treatment regardless of whether the principal has correct beliefs about the returns to the technology, or the distribution of agents' types. While more sophisticated mechanisms could potentially extract some more information (for instance, an agent's beliefs about others' returns) we believe that robustness is an essential property if these mechanisms are to be used in practice.[9]

[6]Note that the percentage of agents rejecting, or opting-in to, treatment is often non-trivial. For example, 45% of the people Dupas and Robinson (2009) opened a savings account for never made a deposit, 72% of the people offered a commitment saving product by Ashraf, Karlan, and Yin (2006) rejected it, and in a study of educational vouchers in Columbia, Angrist, Bettinger, Bloom, King, and Kremer (2002) find that 25% of those randomly denied a voucher were awarded other scholarships, and 10% of those who were offered vouchers declined them.

[7]Even in medical trials with double-blind designs unobserved refusal to accept treatment will still alter the results of the study. For a brief review of RCTs in medicine see Stolberg, Norman, and Trop (2004). Jadad (1998) provides a comprehensive review.

[8]Note that selective trials only allow agents to select themselves probabilistically, i.e. high value agents may select themselves into a group that has higher probability of being treated, but some of them will still end up in the control group.

[9]For an in-depth analysis of robust mechanisms, see Bergemann and Morris (2005) and the references therein.

However, selective trials are not a panacea. Incentive compatibility constraints imply that, in equilibrium, agents with higher values will be assigned to the treatment group with higher probability. This sampling constraint produces over-sampling at the top and under-sampling and the bottom. In an environment with finitely many agents, these sampling constraints adversely affect the statistical power of selective trials. From a theoretical perspective, these losses can be made arbitrarily small by reducing the slope of the probability-of-treatment menu that agents choose from. However, this reduces the agents' incentives to accurately report (or even figure out) their values, which may increase overall noise. A straightforward way to deal with this issue is to increase the number of agents in the experiment.[10]

Our second set of results explore how information production in the first stage affects technology diffusion in the second stage. In that part of the paper we envision the trial as an experiment that the principal runs to convince potential users that the technology is valuable. Whether greater information increases expected usage depends on the shape of the mapping between agents' beliefs and effort, as well as on how agents interpret data.[11] If most agents' beliefs would lead them to use the technology in the absence of new information, a trial that generates noisy information has little upside but might lead to less effort if the trial yields a poor outcome because of noise.[12] Further, agents may hold incorrect beliefs about

[10]Increasing the size of the experiment may be useful regardless of sampling constraints if the experimenter chooses to analyze agents with different values separately. Note that increasing the sample size may be costly, but rarely due to expenses directly related to providing the technology itself. The largest portion of the cost is often incurred measuring the host of observables included as controls in RCTs. These controls are included on the basis that they may be correlated with unobservables such as beliefs about the technology and known privately known idiosyncratic payoff shocks. To the extent that selective trials are validated in the field they may actually allow for larger samples at reduced costs as measurement of these unobservables is built into the experimental protocol, and this may obviate the need for measuring observables.

[11]There is a natural link here to the literature on persuasion games (Milgrom and Roberts, 1986), however, because we assume that the principal cannot hide data from the agents after running the experiment, the strategic concerns in that literature are muted here. Recently, Rayo and Segal (2008) and Kamenica and Gentzkow (2009) consider how the shape of the mapping between agent's beliefs and their actions change incentives for information provision, while Mullainathan, Schwartzstein, and Shleifer (2008) focus on how agents interpret data.

[12]This point is made in a different setting by the industrial organization literature on experience goods

what kind of data the trial would generate in the states of the world where the technology has high returns. This could lead agents to dismiss a technology even though the trial yielded positive results from the principal's perspective. Both of these effects may result in lower usage in the next stage.[13]

While our results suggest that some caution is required, we provide fairly broad sufficient conditions under which information provision has a positive effect on the adoption of useful technologies, and selective trials can be a useful tool to improve adoption.[14] In such settings, by giving the technology to the agents who value it most the principal will generate better information, and increase expected take-up by the agents.[15]

The rest of the paper is structured as follows: Section 2 identifies settings in which our approach is most relevant, Section 3 provides a few examples summarizing the main qualitative points of the paper, Section 4 defines our framework, Section 5 explores how to design randomized experiments to extract more information, Section 6 tackles the problem of inference given data produced by a selective trial, Section 7 explores how the information agents obtain from observing experiments affects their effort decisions, Section 8 discusses various limits to our approach. Appendix A extends the analysis to more general environments. Proofs are contained in Appendix B.

(see, e.g., Bergemann and Välimäki, 2005). Our model differs as the principal and agents may have different beliefs.

[13]This is related to the possibility of inefficient herding in models of social learning, see, e.g. Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992). Note that this refines the natural intuition about the diffusion of new technologies, which implicitly assumes that agents can accumulate good information over time. For example, Suri (2008), states "[A] model where households learn about a technology with positive returns would imply that aggregate adoption rates increase over time." While this may be true in the very long run—see for instance Smith and Sørensen (2000) for theoretical results along these lines—things may be very different in the short-run.

[14]For remarkable experiment that demonstrates how information can successfully affect the behavior of agents, see Dupas (2009a) which studies how information about relative risks of infection affects the sexual behavior of young women in Kenya.

[15]Closely related to the ideas we develop here, Oster and Thornton (2008) find in a policy simulation based on experimental data that targeting the distribution of a new technology (menstrual cups) to users who value the product more would result in greater usage.

2 A Taxonomy of Technologies

While our main focus here is on the use of RCTs in a public health or a development context, our analysis applies to most environments involving decentralized experimentation. For instance, if a firm wants to try a new way to organize production, specific plant managers will have to decide how much effort to put towards implementing it. The firm's CEO is in the same position as the principal in our framework, and must guess the effort exerted by his managers when evaluating the returns to the new technology. Similarly, if a school board wants to experiment with a new program, individual teachers and administrators will have to decide how much effort to expend implementing the program. In what follows we describe the properties of technologies for which our model may be most relevant.

Our model is most applicable to the study of technologies where: 1) there are potentially large private returns, 2) returns depend on the effort expended either learning about or using the technology, 3) usage and effort decisions by agents are imperfectly observable, and 4) agents learn from the outcomes of others. Note that none of these conditions are necessary for the model to apply, or even for the model to be relevant. Rather, these are conditions under which we believe that our approach will be most valuable.

Private vs. Public Returns. Our mechanisms are designed to elicit the agents' private value for the technology. Naturally, this will be most informative when a large part of the returns to the technology are private. For instance, eliciting private valuations would be informative in a study of anti-retroviral drugs where private benefits are large (Thornton, 2008), but less so in a study of deworming drugs where private returns are relatively small (Miguel and Kremer, 2004).

Outcomes Depend on Effort. Eliciting the agents' value for the technology is most useful when the returns to the technology depend significantly on the effort expended on using it properly. If this dependence is small as with, for example, the smallpox vaccine—

a single-dose medication—eliciting valuation may be only marginally useful. If the dependency is large as with, for example, antibiotics or anti-retroviral drugs, eliciting valuation may be more useful. Note that effort can be understood quite broadly and includes effort required to learn how to properly use a technology (see for instance, in the case of fertilizer, Duflo, Kremer, and Robinson, 2008).

Limited Observability of Effort. Eliciting the agents' value for the technology is useful as it gives an indication of how much effort agents are likely to expend. This is only informative if agents' effort is difficult to observe. It is typically the case that at least some dimension of effort is unobservable. Consider for instance the case of fertilizer. One might observe outcomes such as whether a farmer is spreading fertilizer, or how good his understanding of fertilizer use is, but the underlying effort the farmer actually put towards properly using fertilizer is unobservable. Medical treatments taken at home, and contraceptive products are particularly stark examples of technologies where effort is necessary and unobservable.

Learning From Others' Outcomes. Our analysis emphasizes the potential of experiments to promote learning and technology diffusion. This is only useful if agents learn from the outcomes of others (Foster and Rosenzweig, 1995; Kremer and Miguel, 2007; Oster and Thornton, 2008).

Altogether, our framework is likely to be useful when evaluating many (although not all) of the technologies studied in the development or public health literatures, such as water treatment products, fertilizer, antibiotics, or contraceptives.

3 Illustrative Examples

This section uses a simplified version of our general framework to illustrate the main points of the paper. We first explore the informativeness of selective trials versus RCTs, and then consider how the information generated from a trial can increase or decrease future adoption and usage of a product. With respect to information production, we show that when agents are allowed to select into a trial by paying a price, the agents' outcomes will provide a clearer signal of the technology's benefit. This remains true whether agents have heterogeneous beliefs about the technology, heterogeneous costs of effort, or if there are heterogenous private returns to using the technology. With respect to technology adoption, we show that even though a selective trial will provide superior information, in certain situations this may decrease rather than increase expected usage even when the technology provides high returns.

In this section we use the example of a water treatment product for concreteness. However, because our framework applies to a variety of environments, it would be perfectly appropriate to think through the implications in terms of other technologies, such as fertilizer or a drug that must be taken regularly.

3.1 Framework

While the model in this section is simplified, unless specifically noted, the notation and concepts will carry through to the more general framework. There are infinitely many agents indexed by $i \in \mathbb{N}$. For any sequence of random variables $(Z_i)_{i\in\mathbb{N}}$, we define

$$\int Z_i \, \mathrm{d}i \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} Z_i$$

whenever the limit exists. Each agent has a treatment status $\tau_i \in \{0, 1\}$. If agent i is in the treatment group, $\tau_i = 1$, and he is given the water treatment product by the principal. Otherwise $\tau_i = 0$ and the agent is in the control group.

Agent i has a public outcome y_i (e.g. number of non-sick days), that depends on his usage of the water treatment product:

$$y_i = (R + r_i)e_i + \varepsilon_i,$$

where $e_i \in \{0, 1\}$ is agent i's decision of whether or not to put effort into using the product, i.e. adding it to his water, $R \in \{R_L, R_H\}$ is the component of the technology's return that is common to all agents, r_i is agent i's privately known idiosyncratic return to the technology, and ε_i is an idiosyncratic shock with unknown expectation. When the agent is not given the technology $e_i = 0$, so that an agent i in the control group will have an outcome $y_i = \varepsilon_i$. Thus, ε_i is the baseline health outcome that is controlled for using randomization. Given effort e_i, agent i has expected utility

$$u_i(e_i) = \mathbb{E}_i[y_i|e_i] - c_i e_i, \tag{1}$$

where c_i is agent i's cost of effort, which is observationally equivalent to a preference for using (or not) the technology. Both the idiosyncratic return r_i and cost c_i are privately known to each agent i before they decide whether or not to add the product to their water. We assume that the maximum cost of effort is $\bar{c}$, and $R_L < \bar{c} < R_H$.

The expectation in (1) is indexed by i, reflecting the fact that agents have heterogenous beliefs about the common return R. Specifically, each player i has a private belief $\theta_i = \text{Prob}_i(R = R_H)$ that returns are high. Beliefs $(\theta_i)_{i\in\mathbb{N}}$ are distributed according to some distribution F_θ, which need not be known to the principal or the agents. Note that agents do not share a common prior, so others' beliefs are not informative.

The difference in average outcomes between agents who are given the water treatment product and those who are not is

$$\widehat{R} = \frac{1}{\int \tau_i \mathrm{d}i} \int y_i \mathbf{1}_{\tau_i=1} \, \mathrm{d}i - \frac{1}{1 - \int \tau_i \mathrm{d}i} \int y_i \mathbf{1}_{\tau_i=0} \, \mathrm{d}i,$$

which is observable to both the principal and agents. We contrast two ways of running a trial of the water treatment product:

1. An RCT, where agents are randomly assigned to the treatment group with probability π_0 or

2. A selective trial where each agent has the option of paying a price $p \in (0, R_H - \overline{c})$ to receive the product with probability π_0.[16] Agents who do not pay price p do not receive the treatment.

The public outcome is denoted by $\widehat{R}_0$ for the RCT, and by $\widehat{R}_p$ for the selective trial.

The following subsections compare how these protocols produce information, before turning to the question of which protocol is better suited to increase take-up of a new technology that the principal knows to be good.

3.2 Information Production

3.2.1 Heterogeneous Priors

To explore the role of heterogeneous priors in information production, we first consider the case where all agents' costs and returns are common, that is, for all i, $c_i = \overline{c}$ and $r_i = 0$.

In an RCT, when agent i is provided with a water treatment product, she chooses to expend effort using it if and only if $\theta_i R_H + (1 - \theta_i) R_L - \overline{c} \geq 0$, which simplifies to $\theta_i \geq$

[16] Note that there is no need for the probability of treatment in the selective trial to be the same as in an RCT; we use π_0 here for notational simplicity. The experimenter could set the probability of treatment to any $\pi_p \in (0, 1]$.

$\frac{\bar{c}-R_L}{R_H-R_L} \equiv \underline{\theta}$. Hence,

$$\begin{aligned}\widehat{R}_0 &= \frac{1}{\pi_0}\int(R\mathbf{1}_{\theta_i\geq\underline{\theta}} + 0\times\mathbf{1}_{\theta_i<\underline{\theta}} + \varepsilon_i)\mathbf{1}_{\tau_i=1}\mathrm{d}i - \frac{1}{1-\pi_0}\int\varepsilon_i\mathbf{1}_{\tau_i=0}\,\mathrm{d}i \\ &= R\times\mathrm{Prob}(\theta_i\geq\underline{\theta}) = R\times(1-F_\theta(\underline{\theta})).\end{aligned}$$

When the principal knows F_θ, the distribution of agent beliefs, then an RCT will allow her to identify R. However, in most cases the principal does not know F_θ, and $\widehat{R}_0$ only provides a garbled signal of R. If the outcomes y_i of agents in the treatment group aren't particularly good compared to agents in the control group, the principal does not know if this is because returns are low, $R = R_L$, or because few agents put effort into using the product.[17]

In a selective trial, by contrast, all agents who pay the price p are such that they will find it worthwhile to add the product to their water, thus $\widehat{R}_p = R$. An agent chooses to participate in the trial if and only if $\pi_0(\theta_i R_H + (1-\theta_i)R_L - \bar{c}) > p$, which is equivalent to $\theta_i > \frac{p/\pi_0+\bar{c}-R_L}{R_H-R_L} \equiv \overline{\theta} > \underline{\theta}$. As $\overline{\theta} > \underline{\theta}$, all agents who choose to participate in the trial will use the product, given the chance.[18] Thus, $\widehat{R}_p = R$, in spite of aggregate uncertainty over the distribution of beliefs, F_θ.

This example shows that if an underlying selection process is affecting measurable outcomes, then more informative signals can be generated by making this selection process explicit rather than implicit. In the following subsections, we show how this intuition extends in the presence of privately-known heterogeneous costs and returns.

3.2.2 Heterogeneous Costs (or Preferences)

When agents have heterogeneous privately-known costs, $c_i < \bar{c}$, then the signal $\widehat{R}_0$ generated by an RCT is garbled by the heterogeneity of costs and beliefs, but signal $\widehat{R}_p$ from a selective

[17]Note that the same reasoning holds from the perspective of agents.

[18]We implicitly assume that $F_\theta(\overline{\theta}) < 1$, i.e. the price p is not too high (and the probability π_0 is not too low), so that a positive measure of agents participate in the experiment.

trial is still perfectly informative. Agent i chooses to participate if and only if $\pi_0(\theta_i R_H + (1-\theta_i)R_L - c_i) > p$, or equivalently, when $\theta_i > \frac{p/\pi_0 + c_i - R_L}{R_H - R_L} > \frac{c_i - R_L}{R_H - R_L}$. Once again if agent i chooses to pay for the water treatment product, it is optimal for him to put effort into using it, so $\widehat{R}_p = R$.

3.2.3 Heterogeneous Returns

If agents have privately known idiosyncratic returns $r_i \neq 0$ then a selective trial will still be more informative than an RCT, but it will no longer provide a noiseless signal of the common return.

For simplicity, we return to the case where $c_i = \overline{c}$, $\forall i$. Under an RCT agent i chooses to expend effort if and only if $\theta_i \geq \frac{\overline{c} - r_i - R_L}{R_H - R_L} \equiv \underline{\theta}_i$. Hence,

$$\begin{aligned}
\widehat{R}_0 &= \frac{1}{\pi_0}\int ((R + r_i)\mathbf{1}_{\theta_i \geq \underline{\theta}_i} + \varepsilon_i)\mathbf{1}_{\tau_i=1}\mathrm{d}i - \frac{1}{1-\pi_0}\int \varepsilon_i \mathbf{1}_{\tau_i=0}\,\mathrm{d}i \\
&= R \times \mathrm{Prob}(\theta_i \geq \underline{\theta}_i) + \int r_i \mathbf{1}_{\theta_i \geq \underline{\theta}_i}\,\mathrm{d}i
\end{aligned}$$

Estimator $\widehat{R}_0$ is affected by two different sources of noise. The first, $\mathrm{Prob}(\theta_i \geq \underline{\theta}_i)$, reflects uncertainty about what proportion of the treatment group actually uses the product. The second, $\int r_i \mathbf{1}_{\theta_i \geq \underline{\theta}_i}\,\mathrm{d}i$, corresponds to uncertainty about the private returns r_i of agents who do use the product.

If a selective trial is implemented with a small positive price $p = 0^+$, the proportion of agents who participate is $\widehat{\nu}_p = \mathrm{Prob}(\theta_i \geq \underline{\theta}_i)$, and is observable. Note that $\widehat{\nu}_p$ may be much less than 1 as all agents with non-positive value drop from the sample. Thus,

$$\begin{aligned}
\widehat{R}_p &= \frac{1}{\pi_0\widehat{\nu}_p}\int ((R + r_i)\mathbf{1}_{\theta_i \geq \underline{\theta}_i} + \varepsilon_i)\mathbf{1}_{\tau_i=1}\mathrm{d}i - \frac{1}{1-\pi_0\widehat{\nu}_p}\int \varepsilon_i \mathbf{1}_{\tau_i=0}\,\mathrm{d}i \\
&= R + \frac{1}{\widehat{\nu}_p}\int r_i \mathbf{1}_{\theta_i \geq \underline{\theta}_i}\,\mathrm{d}i.
\end{aligned}$$

As $\widehat{R}_0 = \widehat{\nu}_p \times \widehat{R}_p$, knowing $\widehat{\nu}_p$ and $\widehat{R}_p$, is more informative than $\widehat{R}_0$, even though it does not necessarily allow one to infer R. This insight extends to very general settings. Here, offering the water treatment product at some small price extracts greater information, as it identifies the agents who are sufficiently convinced that the technology has high returns to put effort into using it properly.

In a context where outcomes depend on the agents' actions, and hence on the agents' beliefs, greater information may be useful to persuade the agents to put effort in technologies the principal believes to have high returns. In that respect one may consider running informative experiments in the initial phase of a program implementation to drive up adoption. This being said, the next subsection shows that some caution is required, as it is not always the case that better information will lead to greater expected usage of a good technology.

3.3 Technology Diffusion

We remain in the basic framework of Section 3.1, but assume that the principal believes that $R = R_H$ with probability one.[19] The principal is convinced that the technology has high returns and would like to convince the agents to adopt it. The question is whether information can help increase take-up. While more information is always welfare enhancing from the agents' perspective this section shows that this need not be true from the principal's point of view, even though the principal is an altruistic welfare maximizer.

3.3.1 Precision of the Data, and Distribution of Agents' Beliefs

Abstracting away from the process which generates data, suppose that agents can observe a public signal $\widehat{R}$, that takes one of two values, $\widehat{R} \in \{\widehat{R}_H, \widehat{R}_L\}$, and there is common knowledge

[19] As mentioned earlier, in this model differences in beliefs correspond to differences in priors, not differences in information. Therefore, the principal's decisions have no signaling value.

that

$$\text{Prob}(\widehat{R} = \widehat{R}_H | R = R_H) = \text{Prob}(\widehat{R} = \widehat{R}_L | R = R_L) = \pi > 1/2.$$

The principal is generating signal $\widehat{R}$ (for instance by running an experiment) and chooses whether to generate the signal or not. If she chooses to generate signal $\widehat{R}$, the signal is public and cannot be censored. If the principal chooses to give agents access to signal $\widehat{R}$, agent i's ex post beliefs will satisfy:

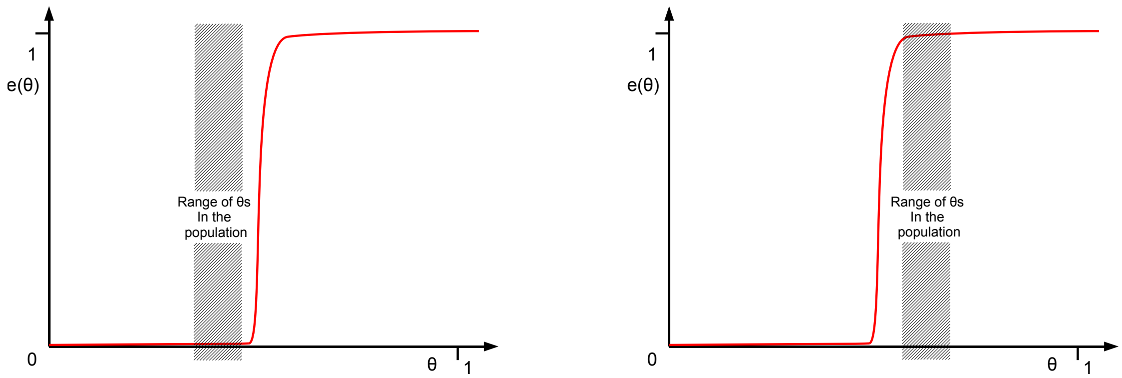
$$\text{Prob}_i(R = R_H | \widehat{R} = \widehat{R}_L) < \theta_i < \text{Prob}_i(R = R_H | \widehat{R} = \widehat{R}_H)$$

where θ_i is agent i's ex ante belief that $R = R_H$. Given a belief θ, we denote by $e(\theta)$ the agent's optimal effort decision.[20]

If $\pi = 1$, the data is perfectly informative. The principal, who is convinced that $R = R_H$, believes the data will be $\widehat{R}_H$ with probability one and from this the agents will conclude that $R = R_H$. In this case it is optimal from the principal's point of view to allow agents to observe the data as it leads them to agree with her. In contrast, when the data is not perfectly informative, i.e. when $1/2 < \pi < 1$, it may mislead the agents in the sense that a bad signal could occur even if the state is good. Whether or not the principal expects data to increase effort depends on the principal's belief about the distribution of $(\theta_i)_{i \in \mathbb{N}}$, and on how the agents' beliefs θ affect their effort choices $e(\theta)$.

As illustrated in Figure 1(a), if most agents have beliefs slightly below the threshold where they put high effort there is little downside to giving agents access to the data (as most of them wouldn't use the product without new information), moreover, there is a significant upside since with high probability the data will be $\widehat{R} = \widehat{R}_H$, which will lead most agents to use the technology. Giving agents access to signal $\widehat{R}$ is optimal.

[20] For this example, we assume that all agents share the same cost for effort, so that optimal effort depends only on beliefs.



(a) Information increases expected effort.

(b) Information decreases expected effort

Figure 1: The effect of noisy information on expected effort, depending on the range of beliefs $(\theta_i)_{i\in\mathbb{N}}$.

In contrast, if as in Figure 1(b) most agents have beliefs slightly above the threshold where they put high effort, then providing access to the data has little upside and a potentially large downside, as a realization of $\widehat{R} = \widehat{R}_L$ would dissuade many agents from expending effort. In this situation, potential noise makes it optimal for the principal not to generate any information.

3.3.2 Disagreements about Data Interpretation

Another environment in which (even very precise) information can reduce expected usage of a good technology is when the principal and agents have different conditional beliefs about what data would be generated conditional on true returns R. That is, it may be that data which is supportive of $R = R_H$, given the principal's conditional beliefs, is actually indicative of $R = R_L$ given agent i's conditional beliefs.

This type of problem may occur if, for instance, the agent has overly high expectations about the returns that would be obtained conditional on $R = R_H$. Consider a setting where signal $\widehat{R}$ takes value in $\{\widehat{R}_L, \widehat{R}_H, \widehat{R}_{HH}\}$ where $\widehat{R}_{HH}$ corresponds to an extremely positive signal. The principal believes that conditional on $R = R_H$ the realization of $\widehat{R}$ will be $\widehat{R}_H$, whereas conditional on $R = R_L$, the realization of $\widehat{R}$ will be $\widehat{R}_L$ or $\widehat{R}_H$ with equal probability. Agents share the belief that conditional o $R = R_L$, $\widehat{R}$ will be $\widehat{R}_L$ or $\widehat{R}_H$ with equal probability. However, agents have overly high expectations conditional on the state being high, i.e. they believe that conditional on R_H the realization of $\widehat{R}$ will be $\widehat{R}_{HH}$ with probability one. Then, as the principal believes that $R = R_H$, she knows that data $\widehat{R}_H$ will occur. This is data is a positive signal given the principal's conditional beliefs, but the agents will conclude that the state of the world really is $R = R_L$. Under such circumstances, the principal will prefer not to generate any information in the first place.

The examples in this section make two points. First, careful experimental design can produce more information about the agents' beliefs and usage. Second, while it seems natural that more information should lead to greater adoption of a good technology, this intuition needs to be applied carefully. In particular, whenever the principal and agents have different beliefs, providing agents with more information is not necessarily welfare improving. We make these points more generally in the next sections.

4 General Framework

We address questions of information production and technology diffusion in a mechanism design framework with two stages. In the first stage we consider the problem of a principal designing experiments to extract more information about the value of a technology. In the second stage we turn to technology adoption and analyze when the principal will want to provide more (or less) information to the agents. We focus on the case where agents are myopic and maximize their current payoffs, utilities are quasilinear, monetary transfers are the only actions that mechanisms can require, and allocations are independent across agents.

We relax these conditions in Appendix A.

Recall that the principal faces infinitely many agents indexed by $i \in \mathbb{N}$. Agents have types t from a type space T, which summarize the agents' beliefs and payoff relevant parameters. As before, types induce an effort decision e, that now takes value in some set $E \subset \mathbb{R}^q$. Returns R to the treatment can now be multidimensional and belong to some set $\mathcal{R} \subset \mathbb{R}^{q'}$. The data $\mathbf{d}$ generated by the agents' outcomes is public and influences both beliefs and effort choices in the second stage.

Types: Each agent i has a type $t_i \in T$, so a profile of types is given by $\mathbf{t} \in T^{\mathbb{N}}$. Types are i.i.d. draws from some distribution $\chi \in \Delta(T)$. Type is decomposed into three attributes $t_i = (\theta_i, \eta_i, F_i)$. Extending the notation of the previous section, $\theta_i \in \Delta(\mathcal{R})$ characterizes agent i's beliefs over returns $R \in \mathcal{R}$. Parameter η_i characterizes the agent's payoffs. In the examples of Section 3, η_i would include both the heterogenous costs to effort c_i, and heterogeneous returns r_i. Finally, $F_i \in \Delta(\Delta(T))$ is the agent's beliefs about the distribution χ of other agents' types.[21] We assume that all measures F_i are absolutely continuous against some measure F and denote by f_i the density of F_i with respect to F. Similarly the principal has some belief $F_P \in \Delta(\Delta(T))$ over the distribution of types, with density f_P against F.

Treatment, effort and utilities: As above, the treatment status of agent i is denoted by $\tau_i \in \{0, 1\}$, with $\tau_i = 1$ indicating the treated group and $\tau_i = 0$ indicating the control group. Assignment mechanisms are formalized below. If agent i is in the treatment group ($\tau_i = 1$), he must make a private effort decision $e_i \in E$. He then obtains an outcome $y_i \in Y$ which depends in part on R. The agent's outcome y_i is drawn from a c.d.f. $F^1_y(R, \eta_i, e_i)$ if $\tau_i = 1$, and drawn from c.d.f. $F^0_y(\eta_i)$ if $\tau_i = 0$. Agent i may be required to make a monetary

[21]Note that we implicitly assume that the profile of types $\mathbf{t}$ is exchangeable and that there is common knowledge that this is the case.

transfer $p_i \in \mathbb{R}$.[22]

Agents have quasilinear utility $v(\eta_i, e_i, y_i) - p_i$, where $e_i = 0$ if $\tau_i = 0$.[23] The indirect utility of an agent with type t_i given assignment (τ_i, p_i) is $V(\theta_i, \eta_i, \tau_i) - p_i$ where,

$$\begin{aligned} V(\theta_i, \eta_i, 1) &= \max_{e \in E} \int_{R \in \mathcal{R}} \int_{y \in Y} v(\eta_i, e, y) \mathrm{d}F_y^1(R, \eta_i, e) \mathrm{d}\theta_i(R) \\ V(\theta_i, \eta_i, 0) &= \int_{y \in Y} v(\eta_i, 0, y) \mathrm{d}F_y^0(\eta_i). \end{aligned}$$

As usual, an agent's preferences are determined up to an affine transformation, so we can normalize $V(\theta_i, \eta_i, 0) = 0$. Thus, $V(\theta_i, \eta_i, 1)$ measures type t_i's value for being assigned to the treatment group. In a slight abuse of notation we denote $V(\theta_i, \eta_i, 1)$ by $V(t_i)$. For simplicity we assume that there exists a known value $V_{\max} > 0$ such that for all $t \in T, V(t) \in [-V_{\max}, V_{\max}]$.

Assignment mechanisms: A mechanism to assign treatment is a pair $G = (M, \mu)$ where M is a set of messages and $\mu : M \to \Delta(\{0,1\} \times \mathbb{R})$ maps individual messages to a probability distribution over assignments τ_i and transfers p_i. Note that in these mechanisms agent i's payment and probability of treatment depend only on his message, but not on messages sent by others.

Beliefs and effort: The optimal effort for type $t_i = (\theta_i, \eta_i, F_i)$ is denoted by $e^*(\theta_i, \eta_i)$ when he is in the treatment group. At the end of the first stage, all agents observe public data $\mathbf{d} = (d_i)_{i \in \mathbb{N}} = (m_i, p_i, \tau_i, y_i)_{i \in \mathbb{N}} \in \mathcal{D}$. Agent i uses this data to form a belief $\theta_i(\mathbf{d}) = \theta_i(\cdot|\mathbf{d})$ over $R \in \mathcal{R}$.[24] We assume common knowledge of rationality, so that each agent believes that

[22] Note that our framework can accommodate observables x_i. An easy way to do this is to include them as part of outcome y_i.

[23] As Appendix A highlights our results extend naturally to general utility functions. This being said quasilinear utility is a good approximation whenever the variation in monetary transfers is small. This variation need only be large enough to provide agents with incentives to figure out their own preferences.

[24] In what follows, θ_i denotes agent i's prior, while $\theta_i(\mathbf{d})$ denotes his posterior after having observed data $\mathbf{d}$.

all other agents play a rationalizable strategy. Thus, $\theta_i(\mathbf{d})$ will depend on i's belief about other agents' types. In the second stage, all agents are given access to the technology and choose effort levels $e^*(\theta_i(\mathbf{d}), \eta_i)$.[25]

We first examine the problem of designing experiments to improve information production, before turning to the question of how to maximize agents' effort in the second stage.

5 Information Production

5.1 The Informativeness of Mechanisms

We use a notion of informativeness for mechanisms that corresponds to informativeness in the sense of Blackwell. Intuitively, a mechanism G is at least as informative as a mechanism G' if the data generated by G' can be simulated using only data generated by G.

The set of rationalizable messages for a player of type t in mechanism G is given by $\mathrm{Rat}^G(t) \subset M$. By extension, we denote by $\mathrm{Rat}^G(\mathbf{t}) = \prod_{i \in \mathbb{N}} \mathrm{Rat}^G(t_i)$ the set of rationalizable message profiles in mechanism G given a profile of types $\mathbf{t}$. We denote by $\mathbf{d}_G(\mathbf{t}, \mathbf{m}, R) : T^{\mathbb{N}} \times M^{\mathbb{N}} \times \mathcal{R} \to \Delta(\mathcal{D})$ the random variable over data profiles generated by mechanism G when the message profile is $\mathbf{m}$ and the state of the world is $(\mathbf{t}, R)$.

Definition 1 (informativeness of mechanisms)**.** *We say that mechanism G is at least as informative as mechanism G', denoted by $G' \preceq G$, if and only if there exists a mapping $h : \mathcal{D} \to \Delta(\mathcal{D})$ such that $\forall \mathbf{t} \in T^{\mathbb{N}}, \forall \mathbf{m} \in Rat^G(\mathbf{t}), \forall R \in \mathcal{R}$,*

$$h(\mathbf{d}_G(\mathbf{t}, \mathbf{m}, R)) \sim \mathbf{d}_{G'}(\mathbf{t}, \mathbf{m}, R).$$

[25] In most settings, this effort decision is multidimensional. For instance, in the case of fertilizer, it is not enough for agents to just expend effort spreading fertilizer. Effort must be expended properly in, for example, choosing the right seeds to go with the fertilizer, or learning how much and when to water the crops. In this case it is natural to think of effort as a vector, where the first component corresponds to picking the right seeds, the second to the right amount of fertilizer, the third to properly applying it, etc.

This is a strong notion of informativeness for mechanisms as it does not depend on the unknown selection mechanism for rationalizable messages. Note that $\preceq$ need not be reflexive when some types have multiple rationalizable messages.

5.2 Benchmark Mechanisms

To anchor the mechanisms we consider to current practice, we propose two benchmark mechanisms. The first, $G_0 = (\emptyset, \mu_0)$, corresponds to the standard randomized controlled trial (RCT) where all agents are assigned to the treatment group with the same probability $\pi_0 \in (0, 1)$. The second, $G_{0^+} = (T, \mu_0)$, is an RCT where in addition agents can make costless arbitrary statements about their type, for instance through a pre-treatment survey. In both mechanisms agents are not required to make any transfers ($p = 0$).

Our first result shows that whenever a mechanism assigns a positive proportion of each type to both the treatment and the control group, it is at least as informative as the benchmark mechanism G_0.

Proposition 1 (full support sampling)**.** *Consider a mechanism $G = (\mu, M)$ with M finite. If there exists $\xi > 0$ such that*

$$\forall m \in \cup_{t \in T} Rat^G(t), \quad \mathbb{E}_\mu(\tau_i | m_i = m) \in [\xi, 1 - \xi],$$

then $G_0 \preceq G$.

This proposition shows that to be as informative as an RCT, it is sufficient that every type have positive probability to end up in the control group and positive probability to end up in the treatment group. While this may suggest that sampling rates do not matter much, it is important to recognize that this result relies on having an infinite number of agents. With a finite number of agents, sampling rates would matter significantly for the power of our estimators. We discuss this point in greater detail in Section 5.3.

Our second result shows that if we are willing to consider mechanisms in *weakly* dominant strategies, then benchmark mechanism G_{0^+} can extract all the information that agents have.

Proposition 2 (full revelation in weakly dominant strategies)**.** *Mechanism G_{0^+} is such that truthful revelation is weakly dominant for each type.*

The intuition for this result is simple: as an agent's inclusion in the treatment group is not dependent on the message they send, sending a message which reveals his type is just as good for the agent as sending any other message. Unfortunately, this means that any other message is weakly dominant as well. Hence, although full revelation is possible here, data generated by G_{0^+} is likely to be unreliable, especially if figuring out one's preferences is costly. Indeed, as Kremer and Miguel (2007) and others have noted, reported beliefs about the returns to a technology are often uncorrelated with willingness to pay. Thus, we turn to implementation in strictly dominant strategies.

5.3 Selective Trials

A selective trial is a mechanism where agents are asked to indicate their value for the technology, and then randomly assigned to the treatment or control group based on their report. As we are concerned with the implementability of selective trials, we focus on mechanisms where each agent has a strictly dominant strategy.[26]

With quasi-linear utilities, describing *a* most informative selective trial is straightforward. Consider the selective trial G^* such that $M = [-1, 1]$, any agent sending message m is assigned to the treatment group with probability $\pi_m = \frac{1+m}{2}$ (independently of the treatment of other agents), and must make a transfer $p(m) = \frac{V_{\max}}{4} m^2$. One can think of agents as having a baseline probability of being in the treatment group equal to 1/2 and deciding by how much they want to deviate from this baseline. An agent with value $V(t)$ chooses

[26] For some of the mechanisms that follow, an agent with value V has a strictly dominant strategy only for almost every value V (with respect to the Lebesgue measure on $\mathbb{R}$), rather that every value.

message m to maximize

$$\pi_m V(t) - p(m) = V(t)\frac{m+1}{2} - V_{\max}\frac{m^2}{4}.$$

This problem is concave in m, and first order conditions yield an optimal message $m^* = V(t)/V_{\max}$. Our first result highlights that no mechanism in strictly dominant strategies can reveal more information than G^*.

Proposition 3 (value revelation in strictly dominant strategies)**.** *Any mechanism G such that every type t has a strictly dominant message is such that $G \preceq G^*$.*

This implies that the only information we can extract from an agent in dominant strategies is his value for being in the treatment group. This provides a useful bound on what one can hope to reveal using mechanisms. Another way to say this is that for any mechanism in strictly dominant strategies, the message m sent by the agent is only a function of the agent's value $V(t)$. Note that the mechanism used to elicit agents' values does not depend on the principal's beliefs about the distribution of agents' types. In that sense G^* and the mechanisms we describe below are robust in the sense of Wilson (1987) or Bergemann and Morris (2005).

While mechanism G^* is relatively simple, letting agents choose from a continuous menu of lotteries may be unattractive. In what follows we restrict attention to mechanisms $G = (M, \mu)$ such that M has cardinal $n < \infty$ and every type t has a strictly dominant action. For simplicity, but without loss of generality, we assume that the set of possible values $\{V(t_i)|t_i \in T\}$ is convex.[27] Also without loss of generality, we focus on mechanisms without redundant messages and assume that every message in M is rationalizable for some type. Finally, given sequences $\pi = (\pi_k)_{k=1,\cdots,n} \in [0,1]^n$ and $\mathbf{p} = (p(k))_{k=1,\cdots,n} \in \mathbb{R}^n$, we denote by

[27] If it's not the case, set equality (2) in Proposition 4 holds as long as each side is intersected with $\{V(t)|t \in T\}$.

$G^{\pi,\mathbf{p}}$ the mechanism such that $M = \{1, 2, \cdots, n\}$ and an agent that sends message k pays cost $p(k)$ and is assigned to the treatment group with probability π_k.

Proposition 4 (a representation result)**.** *Consider a mechanism $G = (M, \mu)$ such that M has cardinal n. There exists a mapping $g : M \to \{1, \cdots, n\}$, as well as numbers $V_0 < V_1 < \cdots < V_n$, $\pi_1 < \cdots < \pi_n$ and a mapping $p : \{1, \cdots, n\} \to \mathbb{R}$ such that for all $m \in M$,*

$$\begin{aligned} \{V(t_i) | Rat^G(t) = m\} &= (V_{g(m)-1}, V_{g(m)}) \\ \mathbb{E}_\mu(\tau_i | m_i = m) &= \pi_{g(m)} \\ \mathbb{E}_\mu(p_i | m_i = m) &= p(g(m)). \end{aligned} \tag{2}$$

Conversely, the mechanism $G^{\pi,\mathbf{p}}$ is such that a type t_i with value $V(t_i)$ will find it strictly optimal to send message j if and only if $V(t_i) \in (V_{k-1}, V_k)$.

Furthermore, given strictly increasing value thresholds $(V_k)_{k=0,\cdots,n}$ and strictly increasing probabilities of treatment $(\pi_k)_{k=1,\cdots,n} \in (0,1)^n$, it is straight-forward for an experimenter to design a mechanism that partitions agents' value according to thresholds $(V_k)_{k=0\cdots n}$. Specifically the sequence of transfers $(p(k))_{k=1,\cdots,n}$ is such that $G^{\pi,\mathbf{p}}$ categorizes values $V(t)$ according to the partition $\{(V_{k-1}, V_k) | k = 1, \cdots, n\}$ if and only if

$$p(k) = p(k-1) + (\pi_k - \pi_{k-1}) V_{k-1}. \tag{3}$$

Note that the sequence of transfers is entirely determined by $p(1)$.

Proposition 4 highlights that general selective trials can be thought of as a simple generalization of RCTs in which the agents' likelihood of being in the treatment group depends on their willingness to make monetary transfers. Beyond this, Proposition 4 has two main implications. The first is that in *any* dominant strategy mechanism, agents can be categorized according to an increasing sequence of thresholds $(V_k)_{k\in\{0,\cdots,n\}}$. The information revealed by the mechanism is which interval (V_k, V_{k+1}) the agent's value belongs to. Furthermore all agents whose value belongs to (V_k, V_{k+1}) are sampled into the treatment group with the same probability π_k and the sampling probabilities $(\pi_k)_{k\in\{1,\cdots,n\}}$ are strictly increasing in k. This implies that agents with high values are over-sampled and those with low-values are under-sampled.

With infinitely many agents, sampling rates do not affect the informativeness of the data, but this would no longer be the case with finitely many agents. In order to maximize statistical power at each point along the distribution of values, it it is often desirable to assign half of the agents with a given value the treatment group and and the other half to the treatment groups. The fact that sampling rates must increase with value means that this cannot be achieved and statistical power will generally not be maximized. As Proposition 5 shows, this can be mitigated to the extent that the sequence $(\pi_k)_{k\in\{1,\cdots,n\}}$ must be strictly increasing, but can in theory be arbitrarily flat.

Proposition 5 (approximately flat sampling)**.** *Pick any $\underline{\rho} < \overline{\rho}$ in $(0,1)$. For any mechanism $G = (M, \mu)$ with M finite, there exists a mechanism $G' = (M, \mu')$ such that $G \preceq G'$, and*

$$\forall m \in M, \quad \mathbb{E}_{\mu'}(\tau_i | m_i = m) \in [\underline{\rho}, \overline{\rho}].$$

Example 1: Mechanism G^* can be generalized to mechanisms G^*_λ such that $M = [-1, 1]$, $\pi_m = \frac{1+\lambda m}{2}$ and $p(m) = \frac{\lambda V_{\max} m^2}{4}$, where λ is the slope of the sampling profile. Note that as λ goes to zero, each agent will be sampled with probability approaching $\frac{1}{2}$, irrespective of the message they send. Sending message $m = V(t)/V_{\max}$ is still a dominant strategy for an agent of type t.

Still the constraint remains that one cannot over-sample low value agents and under-sample high value agents. Furthermore, as Proposition 6 shows, reducing over- and under-sampling to negligible levels, reduces the agents' incentives to send truthful messages. To put this another way, significant over- and under-sampling may be necessary to generate reasonably large incentives for agents to reveal their valuation. To formalize this, define $U(t|m,G) = \pi_m V(t) - p(m)$, the conditional utility of type t when sending message m in mechanism G.

Proposition 6 (incentives and sampling rates)**.** *Consider a mechanism $G = (M, \mu)$ with M finite such that every message is rationalizable for some type and such that $\forall m \in M, \mathbb{E}_\mu(\tau_i | m_i = m) \in [\underline{\rho}, \overline{\rho}]$, then we have that $\forall t \in T$*

$$\max_{m,m' \in M} \quad |U(t|m,G) - U(t|m',G)| \leq 2(\overline{\rho} - \underline{\rho})V_{\max}.$$

In particular, if agents have to learn their own value, providing them with incentives to learn will require the sampling scheme to exhibit sufficient slope.

Example 1 (cont.): Consider an environment in which agents do not know their value for the good but instead have a uniform prior over $[-V_{\max}, V_{\max}]$. Agents can learn their own value by paying a cost k. Mechanism G^*_λ will lead agents to learn their own value if and only if the value of knowing one's type and sending optimal message $m^*(t)$ rather than $m = 0$ is greater than cost k, i.e. if and only if

$$\mathbb{E}_t[U(t|m^*, G^*_\lambda)] \geq k \iff \lambda \geq 12\frac{k}{V_{\max}}.$$

Taken together, Propositions 4, 5 and 6 imply that using selective trials is costly to the extent that it leads to oversampling of high value agents and undersampling of low value agents. However, this cost can be made small if there are many agents or if agents are very responsive to incentives.

5.4 Possible Implementations of Selective Trials

There are several ways to implement selective trials. Proposition 4 suggests an implementation as a menu of lotteries with varying prices and likelihood of treatment. This has the benefit of highlighting the close relationship between selective trials and RCTs. We give an example in which agents are categorized in four bins.

Example 2: Suppose the principal wishes to segment the population into four groups according to the partition $\{(-V_{\max}, 0), (0, \frac{V_{\max}}{4}), (\frac{V_{\max}}{4}, \frac{V_{\max}}{2}), (\frac{V_{\max}}{2}, V_{\max})\}$, that is, she would like to divide agents into those who think the technology will harm them on the on hand, and agents with low, intermediate and high value for the technology on the other hand. The principal wants to make sure that those that think they will be damaged by the technology will not receive it or have to pay to be taken out of the experiment, so she sets $p(1) = 0$ and $\pi = (0, \frac{1}{3}, \frac{1}{2}, \frac{3}{4})$. According to recurrence condition (3), this determines the price schedule and we obtain $\mathbf{p} = (0, 0, \frac{V_{\max}}{24}, \frac{V_{\max}}{6})$.

An other common mechanism for eliciting values in experimental settings is that of Becker, DeGroot, and Marschak (1964) (BDM).[28] In the BDM mechanism the agent sends a message indicating his value. The principal then compares the value conveyed by the agent's message to a pre-selected benchmark value: if the agent's message indicates a value greater than the benchmark, the agent pays the benchmark and receives the technology. If the agent's value is less than the benchmark, no transaction occurs. In a BDM implementation of selective trials, we randomize the benchmark and this randomization is done independently for each agent according to the same distribution. If the message space includes $[-V_{\max}.V_{\max}]$ and $F_{BDM}(\cdot)$ has full support over $[-V_{\max}, V_{\max}]$, it is a strictly dominant strategy for each agent to truthfully reveal their value.

[28] For an overview see Bohm, Lindén, and Sonnegård (1997) and references therein.

Proposition 7 (BDM Implementation)**.** *Every mechanism G in strictly dominant strategies has an equivalent BDM implementation.*

The proof is straight-forward. Given any mechanism G, denote by $\pi_G(V)$ the probability that an agent with value V gets the technology. Note that by Proposition 4, we know that $\pi_G(V)$ must be increasing.[29] Therefore, π_G is a c.d.f. over $\mathbb{R}$. It is immediate that the BDM mechanism with c.d.f. $F_{BDM} = \pi_G$ yields the same data as mechanism G.

Example 2 (cont.): In the case of Example 2, c.d.f. F_{BDM} is given by

$$F_{BDM}(V) = \begin{cases} 0 & if \quad V \leq 0 \\ 1/3 & if \quad V \in (0, V_{\max}/4) \\ 1/2 & if \quad V \in (V_{\max}/4, V_{\max}/2) \\ 3/4 & if \quad V > V_{\max}/2. \end{cases}$$

Each message of the original mechanism is associated with a step in c.d.f. F_{BDM}.

6 Information and Inference

Selective trials extract the maximum amount of information from agents independently of the underlying model generating outcomes. However, using this data to make inferences about underlying parameters necessarily requires a prior over the possible underlying structural models. In this section we consider inference from the principal's perspective. In Section 7 we study how data affects the beliefs and behavior of agents.

[29]This implies that $\pi_G(V)$ has a right-continuous version. We identify the two functions.

6.1 Robust Identification of Marginal Treatment Effects

The mechanisms we describe in Section 5 are robust in the sense that they elicit the agents' valuation for every possible profile of types $\mathbf{t} \in T^{\mathbb{N}}$. In addition, the subsequent randomization of agents in the treatment or control groups allows us to identify treatment effects conditional on an agent's valuation. This corresponds exactly to the marginal treatment effects (MTEs) studied by Heckman and Vytlacil (2005).

MTEs are extremely useful in evaluating the effect of a treatment in different policy contexts, provided the distribution of agents' types does not change. For example, a policy maker might like to use the results of an RCT to understand the effect of partially subsidizing access to a program. Clearly, those who decide to join the program at its subsidized rate are not representative of the general population. Thus, the results of the RCT are not necessarily useful in projecting the impact of improved accessibility on outcomes. In contrast, the MTEs generated by a selective trial can be used to predict the effect of the program for those who are likely to join the program for different levels of subsidization. In the same way that randomized controlled trials provide assumption-free instruments to measure the effect of treatment on the treated, selective trials provide assumption-free *local* instruments that Heckman and Vytlacil (2005) show are necessary to identify MTEs.[30]

MTEs contain the necessary information to perform policy simulations only if the distribution of agents' types is fixed across different policies. This requirement may fail in a number of settings. In particular, to the extent that beliefs (in our notation θs) are malleable and respond to data, the distribution of types in the same population of agents may change from one period to the next. In addition, some policies, for instance advertisement campaigns, can change the agents' preferences (in our notation ηs). In such circumstances, one may be interested in uncovering structural parameter R to predict how policies affecting beliefs or preferences would change the realized returns to the technology.

[30]See Moffitt (2008) for other recent work on the estimation of marginal treatment effects.

6.2 Sufficient Conditions for Full Inference

Given a state R, a distribution of types $\chi \in \Delta(T)$, and an interval of values $\mathcal{V} \subset \mathbb{R}$, let us denote by $\widehat{F}_y^1(R,\chi,\mathcal{V})$ the distribution of outcomes among agents with value $V \in \mathcal{V}$ when assigned to the treatment group. Similarly let us denote by $\widehat{F}_y^0(\chi,\mathcal{V})$ the distribution of outcomes among agents with value $V \in \mathcal{V}$ when assigned to the control group.

Since selective trials identify the agents' willingness to pay for the treatment, they are sufficiently informative to infer state R only if knowing the agents' values allows to disentangle potential confounding effects. This is made more specific in the following definition.

Definition 2 (distinguished states)**.** *We say that two states R and R' in $\mathcal{R}$ are distinguished if and only if for any two distributions of types χ and χ' there exists an interval $\mathcal{V}$ such that*

$$(\widehat{F}_y^1(R,\chi,\mathcal{V}), \widehat{F}_y^0(\chi,\mathcal{V})) \neq (\widehat{F}_y^1(R',\chi',\mathcal{V}), \widehat{F}_y^0(\chi',\mathcal{V})). \tag{4}$$

The purpose of Definition 2 is the following. We want to use data on outcomes to distinguish whether the fundamental state is R or R'. The difficulty is that potentially, the same outcome data may be rationalized different combinations of returns and distributions of types, (R,χ) and (R',χ'). Condition (4) ensures that any two possible rationalizations of the outcome distributions can be distinguished by considering the conditional distribution of outcomes for agents with values in some interval $\mathcal{V}$. The following result is immediate.

Proposition 8. *Selective trials allow to identify the fundamental state R for any realization of the data if and only if any two states R and R' are distinguished.*

While distinguishability is a strong condition, it will be satisfied in a number of relevant environments.

Example 3: Consider for instance the example in Section 3.2, except that we allow R to take value in the entire interval $[R_L, R_H]$ and e to take value in $[0,1]$. As before, $y_i^1 = Re_i + \varepsilon_i$

and $y_i^0 = \varepsilon_i$. If there exists V^+ such that for all types t satisfying $V(t) > V^+$, $e^*(t) = 1$, then any mechanism which identifies agents with value greater than V^+ allows the principal to infer R.

An other interesting case is one where there is uncertainty over the technology's baseline returns, as well as the technology's returns to effort.

Example 4: Imagine that state R takes the form $R = (R^b, R^e) \in \mathbb{R}^2$. Parameter R^b corresponds to baseline returns for being in the treatment group and parameter R^e corresponds to additional returns to effort for those in the treatment group. We have $e \in [0, 1]$ and $y_i^1 = R^e e_i + R^b + \varepsilon_i$ while $y_i^0 = \varepsilon_i$. If there exists V^+ and V^- such that all types with $V(t) > V^+$ choose $e^*(t) = 1$ and all types with $V(t) < V^-$ choose $e^*(t) = 0$, then any mechanism which identifies agents with value below V^- and agents with value above V^+ allows the principal to infer R.

Note that in both of the above examples, inference relies on the knowledge of effort decisions taken by extreme types. This is similar to inference strategies that rely on identification at infinity (Heckman, 1990).

6.3 Incomplete Inference—A Monte-Carlo Study

We now turn to settings in which full inference is not possible and show, by way of simulation, that selective trials can still significantly improve the quality of information generated by an experiment. We consider the model of Section 3 and use a selective trial with three bins, similar to that described in Example 2. Specifically, agents have public outcome:

$$y_i = \tau_i(R + r_i)e_i + \varepsilon_i$$

where, as in Section 3, $R \in \{R_H, R_L\}$ and $e_i \in \{0, 1\}$. Each agent has a private belief $\theta_i = \text{Prob}_i(R = R_H)$, and decides to put effort into using the technology when:

$$e_i = 1 \iff \theta_i R_H + (1 - \theta_i) R_L + r_i - c_i > 0$$

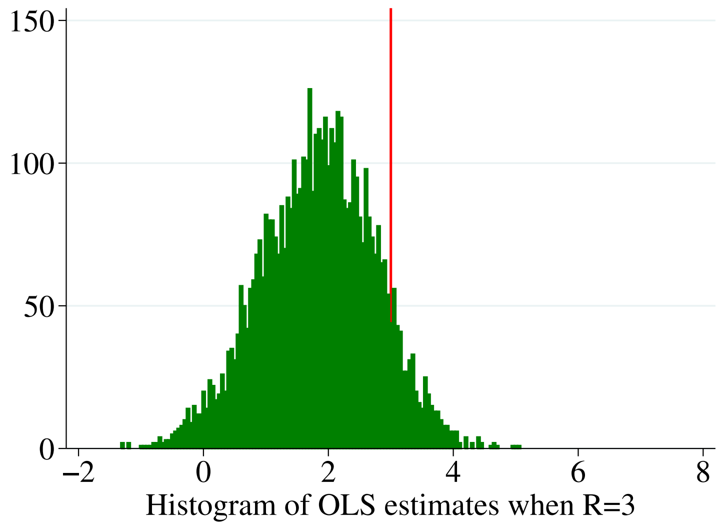
We consider a selective trial where agents can send one of three messages $m_i \in \{1, 2, 3\}$ that are designed such that a message $m_i = 1$ indicates $V_i \in (-\infty, 0]$, $m_i = 2$ indicates $V_i \in (0, 1]$ and $m_i = 3$ indicates $V_i \in (1, \infty)$. Note that all agents who will not expend effort send $m_i = 1$. The treatment probabilities for each message are $\{0.35, 0.5, 0.65\}$, respectively. The principal observes $\{m_i, \tau_i, y_i\}_{i \in N}$, and makes inferences using this data.

For each simulated trial, we have N agents, each with private belief drawn from $\theta_i \sim Beta(1 + \lambda, 1 + (1 - \lambda))$ where $\lambda \sim U[0, 1]$ is an unknown parameter. Uncertainty over λ corresponds to aggregate uncertainty over the distribution of beliefs. Each agent has idiosyncratic returns $r_i \sim N(0, \sigma)$ where $\sigma \sim U[0, \overline{\sigma}]$ is also an unknown parameter, and $\overline{\sigma}$ is known. Uncertainty over σ corresponds to aggregate uncertainty over the distribution of idiosyncratic returns. Altogether, the fact that there is aggregate uncertainty over both idiosyncratic returns and beliefs means that if a population has a large amount of agents with high values, it is hard to determine whether this is because many agents have optimistic beliefs or whether because many agents have high idiosyncratic returns. For this reason, perfect inference is impossible. Finally, $\varepsilon_i \sim N(\mu_\varepsilon, \sigma_\varepsilon)$ where μ_ε and σ_ε are known parameters.

Throughout, we set $R_H = 3$, $R_L = 0.5$, and $c_i = 2$. Additionally, $\overline{\sigma} = 3$, $\mu_\varepsilon = 1$ and $\sigma_\varepsilon = 5$. Finally, we set $\text{Prob}(R = R_H) = 0.5$.

6.3.1 Conditional and Unconditional Treatment Effects

We first use data produced by our simulations to estimate the returns of the technology conditioning on the declared value of agents. Figures 2 and 3 present the results of OLS regressions from 10,000 experiments, each with 200 agents. Note that the unconditional treatment effects estimated using RCTs and displayed in Figure 2 systematically underestimate the true returns to the technology as approximately 56% of the agents do not expend effort.



Notes: Results of 10,000 trials. Data from each trial is used to estimate $y_i = \alpha + \beta\tau_i + \varepsilon_i$, and histogram presents estimates of β for each trial where $R = R_H$. Vertical line indicates R_H.

Figure 2: RCTs only allow the estimation of a single treatment effect, which is systematically biased downwards conditional on $R = R_H$.

Selective trials allow the estimation of the treatment effect conditional on the message sent by the agents, as shown in Figure 3. These estimates may be useful in determining which agents should receive a subsidy for, and which should receive education about, the technology.

The first panel of Figure 3 shows the treatment effect for agents who sent $m_i = 1$. These agents will not put in any effort, and their returns to treatment are on average estimated to be zero. The second panel shows OLS estimates for agents who sent $m_i = 2$ ($V_i \in (0, 1]$), and the third panel shows agents with $m_i = 3$ ($V_i \in (1, \infty)$). The second and third panel are thus the treatment effects conditional on the agents value being between zero and one, and being greater than one. Note that estimates are quite a bit higher when $m_i = 3$. This is to be expected as $\mathbb{E}[r_i|m_i = 3] > \mathbb{E}[r_i|m_i = 2] > 0$. Indeed, agents who send messages indicating they have higher values have, on average, higher idiosyncratic returns. Note also that the variance of estimates is larger in a selective trial because fewer agents are in each group.

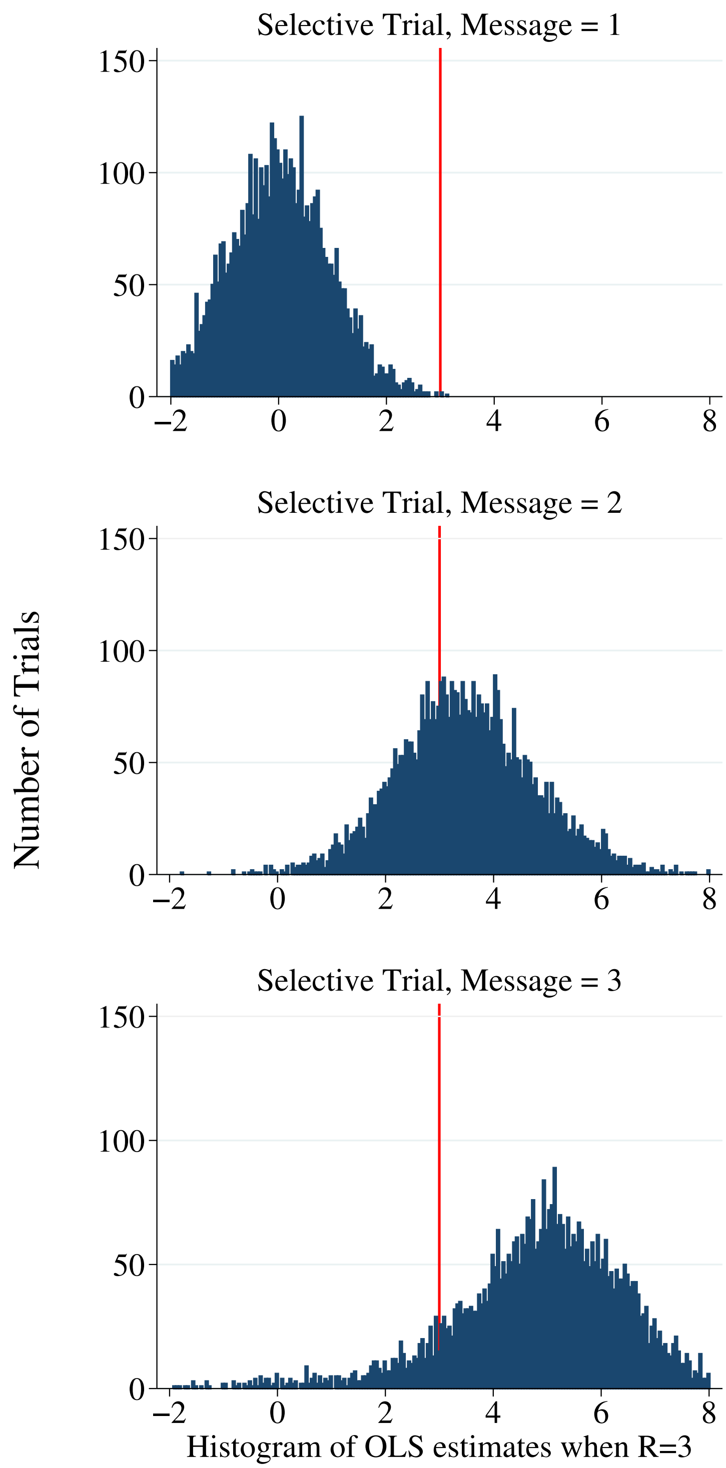
To draw inference from this data while controlling for possible selection bias, i.e. the fact that $\mathbb{E}[r_i|m_i = 3] > \mathbb{E}[r_i|m_i = 2] > 0$, we turn to Bayesian inference.

6.3.2 Bayesian Policy Evaluation

The principal wants to subsidize a technology, and possibly launch an information campaign to encourage take-up if $R = R_H$, but not if $R = R_L$. We use Bayesian inference, and assume that the principal will subsidize the technology if the posterior probability that $R = R_H$ given the data from the trial is greater than 0.5. Throughout this section, the principal knows the data generating process, whether that data is produced by an RCT or a selective trial. Due to computational constraints we use only $N = 100$ agents per trial.

As usual, the principal can make two types of error: a false positive occurs when the principal subsidizes the trial even though $R = R_L$, and a false negative occurs when she fails to subsidize the technology when $R = R_H$. The panels of Figure 4 show how the improvement in inference due to using a selective trial rather than an RCT varies with the percent of agents who do not expend effort ($e_i = 0$).[31] For our limited simulation, selective

[31] The code used to produce the simulations, and documentation on Bayesian inference from this model is available from the authors on request.



<u>Notes:</u> Results of 10,000 trials. Estimates of interaction effects between message and treatment status presented in histograms. Vertical line indicates R_H.

Figure 3: Selective Trials allow the estimation of the returns to technology, conditional on reported values.

trials consistently reduce total errors by about 50%. Note that as fewer agents expend effort the selective trial also samples fewer agents, which results in a slightly smaller improvement in total errors.

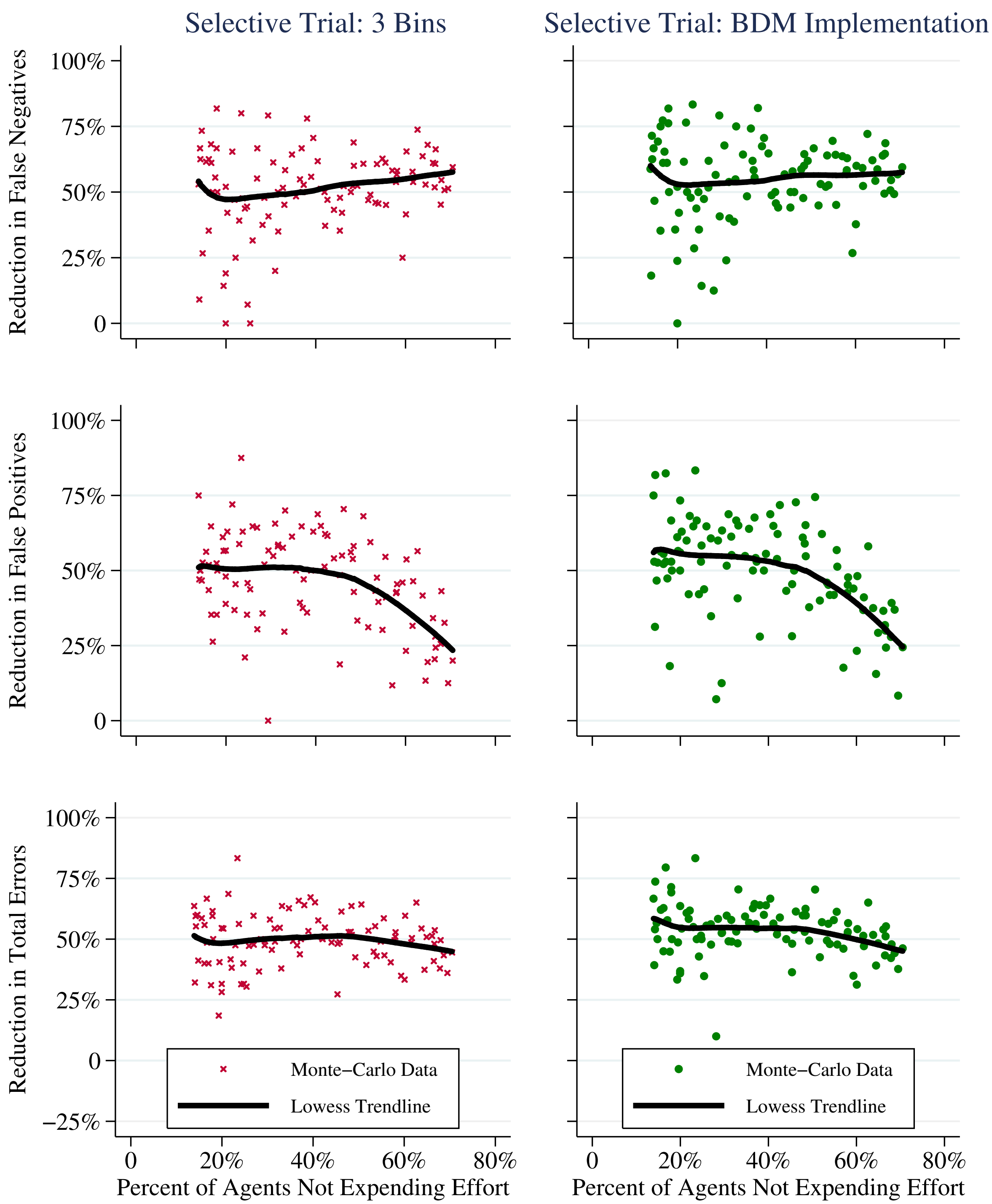
The second column of Figure 4 shows the corresponding error rates when the principal observes data generated by a maximally informative selective trial that identifies each agent's exact valuation. The improvement over the selective trial with three bins is small as much of the informational gains come from knowing which agents will put in effort ($e_i = 1$) and which will not.

Altogether, this section shows that inference depends critically on the principal's beliefs about the way in which data is generated. In the next section we extend this analysis and consider what agents will infer from data. We also examine the ways in which differences in the principal's and agents' beliefs influence the principal's incentives to setup informative experiments for the benefit of the agents.

7 Information, Persuasion and Technology Diffusion

In this section we take the perspective of a principal that knows the technology has high returns and thus would like agents to adopt it. To the extent that the agents' effort decisions depend on their beliefs and that their beliefs respond to information, providing agents with relevant data may improve take-up. A reasonable and salient way to provide information is to run public experimental trials for the benefit of the agents. We focus on a situation where once such a public experiment is setup the principal cannot suppress the data, even if it happens to make the technology look bad. This approach is similar to that of Rayo and Segal (2008) and Kamenica and Gentzkow (2009).[32] Our model differs as we consider

[32] In particular, they also consider a framework in which the principal can commit not to interfere with realized data.



Notes: Each datapoint represents the results of 400 trials with 100 agents each. Reduction in errors is calculated as: $\frac{\text{errors(RCT)}-\text{errors(ST)}}{\text{errors(RCT)}}$

Figure 4: Selective trials consistently reduce inference errors.

an environment where the principal has limited information about the agents' types. In particular we allow for agents to have unknown heterogenous beliefs. As a consequence we do not try to fine tune the optimal information revelation mechanism to the principal's beliefs. Rather, we characterize properties of the environment which ensure that optimal mechanisms take a simple form: either reveal all possible information, or reveal as little information as possible.

We focus exclusively on dominance-solvable mechanisms G. Thus, under common knowledge of rationality, there is a unique mapping from types to messages. This means that a belief $f_i(\mathbf{t})$ over types $\mathbf{t} \in T^{\mathbb{N}}$ induces a unique rationalizable belief over data profiles $\mathbf{d}$. In what follows, we denote by $\mathbf{d}_G$ the data generated by mechanism G and denote by $f_i(\mathbf{d}_G)$ agent i's beliefs over data generated by mechanism G. We use similar notation to denote conditional distributions.

7.1 The Common Prior Case

When the principal and the agents have a common prior over R and $\mathbf{t}$, then it is obvious that providing more information in the first stage improves the welfare of agents in the second stage. This occurs because the principal can trust agents to "do the right thing" with the data (Van den Steen, 2005).

As Section 3.3 highlighted, this is not true anymore when the principal and agent have different beliefs. Indeed, while data will help an agent make better decisions from his perspective, this need not be the case anymore from the principal's perspective. As in many environments with heterogeneous beliefs the principal has incentives for paternalism (see Sandroni and Squintani (2007) for a related point in the context of insurance).

7.2 Data and Beliefs when Priors are Heterogeneous

We first consider when data will move the agents' beliefs closer to those of the principal, before introducing the principal's specific optimization problem in the next subsection. We focus on the case where there is a set of states $\mathcal{R}_H \subset \mathcal{R}$ where the technology can be thought of as having high returns and $\mathcal{R}_L \equiv \mathcal{R} \setminus \mathcal{R}_H$ is the set of states where the technology has low returns. The principal believes $R \in \mathcal{R}_H$ with probability 1. Given data $\mathbf{d}$, agent i's conditional belief that $R \in \mathcal{R}_H$ is:

$$\theta_i(\mathcal{R}_H|\mathbf{d}) = \frac{\theta_i(\mathcal{R}_H)f_i(\mathbf{d}|\mathcal{R}_H)}{\theta_i(\mathcal{R}_H)f_i(\mathbf{d}|\mathcal{R}_H) + \theta_i(\mathcal{R}_L)f_i(\mathbf{d}|\mathcal{R}_L)}.$$

The question is whether data will move belief $\theta_i(\mathcal{R}_H|\mathbf{d})$ closer to $\theta_P(\mathcal{R}_H) = 1$. The change in agent i's belief is given by,

$$\Delta_i(\mathbf{d}) \equiv \theta_i(\mathcal{R}_H|\mathbf{d}) - \theta_i(\mathcal{R}_H) = \theta_i(\mathcal{R}_H)\theta_i(\mathcal{R}_L)\frac{f_i(\mathbf{d}|\mathcal{R}_H) - f_i(\mathbf{d}|\mathcal{R}_L)}{\theta_i(\mathcal{R}_H)f_i(\mathbf{d}|\mathcal{R}_H) + \theta_i(\mathcal{R}_L)f_i(\mathbf{d}|\mathcal{R}_L)}.$$

Upon receiving data $\mathbf{d}$, agent i will increase or decrease his belief that $R \in \mathcal{R}_H$ depending on whether $f_i(\mathbf{d}|\mathcal{R}_H) - f_i(\mathbf{d}|\mathcal{R}_L)$ is positive or not. Note that from the perspective of agent i, $\Delta_i(\mathbf{d})$ has zero expectation, that is, agent i does not expect her beliefs to change in a systematic way. In contrast, as the principal believes that $R \in \mathcal{R}_H$, it may be that $\mathbb{E}_P[\Delta_i(\mathbf{d})] \neq 0$. Intuitively, it is likely that from the principal's perspective $\Delta_i(\mathbf{d})$ will be biased upwards, i.e. giving agents data should lead them to agree with her. However this need not be the case when the principal and agent i hold different beliefs about what data $\mathbf{d}$ will be generated conditional on state R, i.e. when agents and the principal disagree about how to interpret the same data.

We provide conditions which ensure that data will on average push the agents' beliefs closer (or further) from those of the principal.

Definition 3 (weakly optimistic conditional beliefs)**.** *The principal has weakly optimistic conditional beliefs over the data generated by G if and only if for any type t_i and event E,*

$$f_P(\mathbf{d}_G|\mathcal{R}_H, t_i, E) \geq f_i(\mathbf{d}_G|\mathcal{R}_H, E) \iff f_i(\mathbf{d}_G|\mathcal{R}_H, E) \geq f_i(\mathbf{d}_G|\mathcal{R}_L, E). \tag{5}$$

When the principal has weakly optimistic conditional beliefs, she puts weakly higher probability than agent i (conditional on $\mathcal{R}_H$) on data $\mathbf{d}_G$ such that $\Delta_i(\mathbf{d}_G) \geq 0$, i.e. on data that increases agent i's belief that $R \in \mathcal{R}_H$. A particular case of weakly optimistic conditional beliefs is that of *common conditional beliefs*, where $f_P(\mathbf{t}|R) = f_i(\mathbf{t}|R)$, i.e. conditional on state R, agent i and the principal have the same beliefs over the distribution of types in the population, and hence they have the same beliefs about what data $\mathbf{d}_G$ will be generated. Note that this does not preclude heterogenous beliefs over state R. Weakly optimistic beliefs stand in contrast with strongly pessimistic beliefs.

Definition 4 (strongly pessimistic conditional beliefs)**.** *The principal has strongly pessimistic conditional beliefs for the data generated by G if and only if for any type t_i and event E,*

$$f_P(\mathbf{d}_G|\mathcal{R}_H, t_i, E) \geq f_i(\mathbf{d}_G|\mathcal{R}_L, E) \iff f_i(\mathbf{d}_G|\mathcal{R}_H, E) \leq f_i(\mathbf{d}_G|\mathcal{R}_L, E). \tag{6}$$

Condition (6) implies that the principal puts higher probability than agent i (conditional on $\mathcal{R}_L$) on data $\mathbf{d}_G$ such that $\Delta_i(\mathbf{d}_G) \leq 0$, i.e. on data that decreases agent i's belief that $R \in \mathcal{R}_H$. As discussed in Section 3.3.2, a natural environment in which the principal will have strongly pessimistic conditional beliefs is one in which the agents have overly high expectations about the outcomes of treatment conditional on $R \in \mathcal{R}_H$. Imagine for instance that the agent believes that conditional on $R \in \mathcal{R}_H$ the data will look extremely good, to the point that the principal expects such data to have zero probability. Then the principal knows that whenever data has positive conditional probability from her perspective, it is disappointing for the agents. This implies that the principal holds strongly pessimistic conditional beliefs.

As the following proposition shows, Definitions 3 and 4 allow us to characterize when providing agents with more informative data will move their beliefs closer (or further) from those of the principal.

Proposition 9 (biased posteriors)**.** *Consider dominance solvable mechanisms G and G' such that $G' \preceq G$. Whenever the principal holds weakly optimistic conditional beliefs over $\mathbf{d}_G$, then $\mathbb{E}_P[\Delta_i(\mathbf{d}_G)] \geq \mathbb{E}_P[\Delta_i(\mathbf{d}_{G'})]$.*

Whenever the principal holds strongly pessimistic conditional beliefs over $\mathbf{d}_G$, then $\mathbb{E}_P[\Delta_i(\mathbf{d}_G)] \leq \mathbb{E}_P[\Delta_i(\mathbf{d}_{G'})]$.

In words, from the perspective of a principal with weakly optimistic conditional beliefs, giving agents information generates a positive bias in agents' beliefs, whereas, from the perspective of a principal with strongly pessimistic conditional beliefs information generates a negative bias in agents' beliefs.[33]

Intuitively, if the principal expects that information will lead agents' beliefs to converge towards hers, information should improve welfare from the principal's perspective. The last remaining issue is that while information may improve the agents' beliefs on average, it need not improve them for sure when information is noisy. We explore this trade-off in the following section.

[33] If G is strictly more informative than G', in the sense that the additional data extracted by G can change the agent's posterior beliefs, then the inequalities in Proposition 9 are strict.

7.3 Bias, Variance, and the Optimality of Information Provision

In this section we are interested in how data $\mathbf{d}$ affects the agents' effort decisions $e^*(\theta_i(\cdot|\mathbf{d}), \eta_i)$. We consider an environment in which $e \in \mathbb{R}$ and the principal wants to maximize

$$W = \mathbb{E}_P \int e^*(\theta_i(\cdot|\mathbf{d}), \eta_i) \mathrm{d}i. \tag{7}$$

As discussed in the previous section, $\theta_i(\cdot|\mathbf{d})$ is a random variable centered on θ_i from the perspective of the agent but may have some bias from the perspective of the principal.

Given (7), $e^*(\theta_i(\cdot|\mathbf{d}), \eta_i)$ can be seen as the principal's utility over agent i's beliefs. In this context, θ_i is a sure outcome, and $\theta_i(\cdot|\mathbf{d})$ is a lottery over agent i's beliefs. Therefore, the principal's choice to provide information or not can be seen as a choice problem between a sure outcome and a biased lottery. Given this interpretation, it is intuitive that the principal's optimal decision will depend both on the slope and the convexity of mapping e^* with respect to $\theta_i(\cdot|\mathbf{d})$. Distribution $\theta_i(\cdot|\mathbf{d}) \in \Delta(\mathcal{R})$ is potentially a high dimensional object. For the purpose of tractability, we focus on the case where $\mathcal{R} = \{R_L, R_H\}$, so that beliefs $\theta_i \in \Delta(\mathcal{R})$ are unidimensional. We abuse notation and define $\theta_i \equiv \theta_i(R = R_H)$, player i's belief that $R = R_H$. We also define $\theta_i(\mathbf{d}_G) \equiv \theta_i(R = R_H|\mathbf{d}_G)$. Finally, we assume that $\frac{\partial e^*}{\partial \theta} \geq 0$ so that agents with higher beliefs put higher effort, and the principal wants to encourage high beliefs.

To increase the smoothness of the mapping between beliefs and effort, we average effort over different types. To do so, we decompose payoff shock η_i as an aggregate and an idiosyncratic shock, i.e. $\eta_i = (\eta_i^{||}, \eta_i^{\perp})$, where $\eta_i^{||}$ is correlated among agents, and hence is correlated to the data, while $\eta_i^{\perp}$ is independent of all other variables, including the aggregate data. Given $t_i = (\theta_i, \eta_i^{||}, \eta_i^{\perp}, F_i)$, we define

$$\phi(\theta, t_i) = \int_{\eta_i^{\perp}} e^*(\theta, \eta_i) dF_P(\eta_i^{\perp}|\theta_i, \eta_i^{||}, F_i).$$

Note that $\frac{\partial e^*}{\partial \theta} \geq 0$ implies $\frac{\partial \phi}{\partial \theta} \geq 0$. The principal is maximizing $W = \mathbb{E}_P[\phi(\theta_i(\mathbf{d}), t_i)]$. The following example demonstrates the transformation from e^* to ϕ.

Example 5: Recall the example of Section 3. Assume that $u_i(e_i) = \mathbb{E}_i[y_i|e_i] - c_i e_i$, with c_i uniformly distributed over $[0, \bar{c}]$, independent of other data, $R_L < \bar{c} < R_H$ and $e_i \in \{0, 1\}$. In this case, $\eta_i^{\perp} = c_i$. We have $e^*(\theta, c_i) = \mathbf{1}_{\theta \geq \frac{c_i - R_L}{R_H - R_L}}$, which is discontinuous in θ. In contrast, $\phi(\theta, t_i) = [\theta(R_H - R_L) + R_L]/\bar{c}$ is linear in θ.

We now provide sufficient conditions under which more informative mechanisms generate greater expected effort from the agents.

Proposition 10 (sufficient conditions for informativeness to be optimal)**.** *Consider a dominance solvable mechanism G. Assume that for any data $\mathbf{d}_G$, $\theta_i(\mathbf{d}_G) \in [\underline{\theta}, \overline{\theta}]$. If the principal has weakly optimistic conditional beliefs and,*

$$\inf_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial \phi(\cdot, t_i)}{\partial \theta} + \frac{1}{2} \inf_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial^2 \phi(\cdot, t_i)}{\partial \theta^2} \geq 0, \tag{8}$$

then for all dominance solvable mechanisms $G' \preceq G$, we have $W(G') \leq W(G)$.

As ϕ is increasing in θ, a sufficient condition for more information to increase the principal's welfare is for ϕ to be convex. This is a property that is frequently encountered in the industrial organization literature on experience goods. A similar condition appears in Kamenica and Gentzkow (2009). Our results differ as our model allows for heterogeneous priors and thus, the agents' posteriors may be biased from the principal's perspective.

Given Proposition 9, the principal's choice over whether or not to provide information can be framed as a choice between a biased lottery and a sure prize. Therefore it is not surprising that (8) involves a trade-off between the first and second derivative of ϕ. Whenever Proposition 10 applies, information will increase expected effort independently of whether agent i starts with a high or low belief that $R = R_H$, and independently of how informative mechanisms G and G' are.

It is important to note that Proposition 10 holds under the condition that the principal has weakly optimistic conditional beliefs. When the principal has strongly pessimistic beliefs, the agents' posteriors have a negative bias from the principal's perspective, $\mathbb{E}_P[\theta_i(\mathbf{d})] \leq \theta_i$. If this is the case, the next proposition shows that under broad circumstances it will be in the principal's interest to provide as little information as possible.

Proposition 11 (information provision under pessimistic beliefs)**.** *Consider a dominance solvable mechanism G. Assume that for any data $\mathbf{d}_G$, $\theta_i(\mathbf{d}_G) \in [\underline{\theta}, \overline{\theta}]$. If the principal has strongly pessimistic conditional beliefs and,*

$$- \inf_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial \phi(\cdot, t_i)}{\partial \theta} + \frac{1}{2} \sup_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial^2 \phi(\cdot, t_i)}{\partial \theta^2} \leq 0 \tag{9}$$

then for all dominance solvable mechanisms $G' \preceq G$, we have $W(G') \geq W(G)$.

In particular, when Proposition 11 applies, it will be in the principal's interest not to run any experiment. It is tempting to interpret situations where the principal holds strongly pessimistic beliefs as situations where the agents are systematically misinterpreting information. If this is the case, the principal may be able to frame (and simplify) the information in a way that the agents will understand better. In that sense, when information is used to persuade, it may be worthwhile to sacrifice precision for the benefit of clarity.

Example 5 (cont.): Consider ϕ from the previous example. Then, $\frac{\partial \phi}{\partial \theta} = (R_H - R_L)/\overline{c} \geq 0$ and $\frac{\partial^2 \phi}{\partial \theta^2} = 0$. This implies that both (8) and (9) are satisfied. The principal cares only about the expected beliefs of the agents. This implies that whether or not the principal should provide information depends on whether the principal has weakly optimistic or strongly pessimistic conditional beliefs. In the focal case where principal and agents have common conditional beliefs, that is $f_P(\mathbf{d}_G|R,E) = f_i(\mathbf{d}_G|R,E)$, the principal will find it advantageous to run the most informative experiment. If on the other hand, the principal believes that agents are going to be systematically disappointed by data, she should not run any experiment.

8 Discussion

Our mechanism design approach suggests that adding a simple selection mechanism to standard RCTs can increase their informativeness. By eliciting agents' private value of treatment, selective trials help control for unobservable effort decisions that reduce the informativeness of outcomes. Further, under fairly broad conditions, providing agents with better information will increase agents' use of valuable technologies. Thus, selective trials could be used as a way to promote adoption.[34]

Although our approach emphasizes both generality and robustness, it is important to recognize its limitations. First, we focus entirely on eliciting the agents' private value for the technology. This information is useful in settings where the private returns are potentially large, but less useful when the technology's returns are mostly public. More sophisticated mechanisms, such as mechanisms where assignment is not independent across agents, could potentially elicit additional information on agents' beliefs about the public returns of the technology. This is beyond the scope of this paper, but is an attractive avenue for future research.

Second, because selective trials are somewhat more sophisticated than RCTs, they might increase the artificiality of the experiment. This may make experimental disturbances such as the Hawthorne effect more pronounced. The John Henry effect, where agents in the control group feel compelled to compete with agents in the treatment group, may be a particular

[34] This is also suggested by Oster and Thornton (2008).

concern as in a selective trial agents make an explicit choice between lotteries, and some agents who signal a high valuation may be assigned to the control group. It is difficult to anticipate the prevalence of these problems, but we believe they may be minimized by careful experimental design.

We draw attention to several possible options for dealing with experimental disturbances that fit within our model and may be of practical use. First, agents can be dealt with on an individual basis and need not know the assignment of others in the first stage of the experiment. Second, agents in the control group can be given a consolation prize which they may, in fact, prefer to the treatment. This would help reduce the feeling that those who got the treatment are somehow winners and that the losers must catch up with them. Third, our analysis extends naturally to settings where lotteries are associated with different inconvenience costs rather than different prices. For example, agents who want the treatment with high probability may be required to perform boring tasks, sit through long information sessions, take tests, or come back to learn their treatment status. Fourth and finally, the choices that agents make need not be framed as lotteries, in the sense of gambling. For instance, in a BDM implementation, agents only have to pay if they are assigned to the treatment group, although the amount they do pay depends on the message they send.[35] Altogether, we want to emphasize that the mechanisms we describe in this paper have many degrees of freedom and can be adapted to a variety of environments.

A third potential limitation of our approach is that we use a Bayesian model of learning to study the effect of information on future usage, which suggests that we require high levels of rationality on the part of players. However, because we allow the agents to have

[35] Each of these suggestions may have its own drawbacks. Agents that score high on the test will be different from agents that score low. Information sessions that increase knowledge about how to use the technology may be a benefit, rather than a cost. Telling agents assigned to the control group that the treatment is no longer available may make them feel cheated. This may be interpreted as deception and could significantly affect the success of future studies; see Ortmann and Hertwig (2002). We bring these options up not as tried and true methods, but as possibilities practitioners may wish to consider.

heterogenous priors, and in particular hold beliefs that are systematically different from those of the principal, we believe that our approach is flexible enough to accommodate many boundedly rational models of learning. For example, our model can accommodate agents that do not take selection bias into account when they update. Imperfect updating of this form seems to be at the basis of the winner's curse, and appears in many behavioral models that try to improve on the standard Bayesian framework (see e.g., Eyster and Rabin, 2005; Crawford and Iriberri, 2007). In that setting, oversampling high value agents would help promote adoption as other agents do not take into account the fact that the treatment group selected itself on the basis of high anticipated return. Conversely, our model can accommodate agents that systematically overestimate others' effort in using the technology, which would lead them to decrease their estimate of the technology's value when presented with others' outcomes. In such a setting, improving both the quality and the clarity of the data is essential to insure that providing information does not reduce usage of good technologies.

A Extensions

A.1 Non-quasilinear preferences

Section 4 makes the assumption that agents have quasilinear utilities and that mechanisms $G = (M, \mu)$ map messages $m \in M$ to assignments over treatment $\tau \in \{0, 1\}$ and monetary transfers $p \in \mathbb{R}$. In this section we assume that instead of monetary transfers p, the agent is assigned some action $a \in A$, which can either be thought of as an actual action (e.g. wait in line) or as a good (e.g. a pound of sugar, a pack of cigarettes...). For simplicity, we focus on the case where A is finite. A mechanism $G = (M, \mu)$ is now characterized by a message space M and an assignment function $\mu : M \rightarrow \Delta(\{0, 1\} \times A)$.

The principal's effort decision $e \in E$, the possible returns to treatment $R \in \mathcal{R}$, the agents' types $t = (\theta_i, \eta_i, F_i) \in T$, and the outcome distributions $F_y^1(R, \eta, e)$ and $F_y^0(\eta)$ are the same as in Section 4. The preferences of agent i are now represented by a general utility function $u(\eta_i, e_i, y_i, a_i)$.

We denote by $U(\theta_i, \eta_i, \tau_i, a_i)$ the indirect utility of type t_i when her assignment is (τ_i, a_i). Depending on whether or not agent i is assigned to the treatment group, we have,

$$\begin{aligned} U(\theta_i, \eta_i, 1, a_i) &= \max_{e \in E} \int_{R \in \mathcal{R}} \int_{y \in Y} u(\eta_i, e, a_i, y) \mathrm{d}F_y^1(R, \eta_i, e) \mathrm{d}\theta_i(R) \\ U(\theta_i, \eta_i, 0, a_i) &= \int_{y \in Y} u(\eta_i, 0, a_i, y) \mathrm{d}F_y^0(\eta_i). \end{aligned}$$

Definition 5 (normalized utilities)**.** *Denote $\mathcal{U}_0$ the set of normalized utility functions defined by*

$$\mathcal{U}_0 = \{U : \{0,1\} \times A \to [0,1] \mid \min_{(\tau,a) \in \{0,1\} \times A} U(\tau, a) = 0 \textit{ and } \max_{(\tau,a) \in \{0,1\} \times A} U(\tau, a) = 1\}.$$

For every type $t_i = (\theta_i, \eta_i, F_i)$, type t_i's preferences $U(\theta_i, \eta_i, \tau, a)$ over $(\tau, a) \in \{0,1\} \times A$ are uniquely represented by an element of $\mathcal{U}_0$. We normalize all utilities in this way from now on. The data generated by mechanism $G = (M, \mu)$ is a sequence $\mathbf{d} = (m_i, a_i, \tau_i, y_i)_{i \in \mathbb{N}}$. Our notion of informativeness of mechanisms remains that of Definition 1. We denote $U(t)$ the utility function in $\mathcal{U}_0$ that represents the indirect preferences of an agent with type t.

Utility $U(t)$ is the only information that can be elicited from an agent with type t by using mechanisms in dominant strategies, i.e. in any mechanism $G = (M, \mu)$, the best response $m(t)$ of an agent with type t is really a function $m(U(t))$ of type t's indirect utility.The results of Sections 6 and 7 are unchanged provided that we replace $V(t)$ by $U(t)$. We now outline how, as in Section 5, it is possible to elicit the utility $U(t)$ of each type. We only describe one mechanism which achieves that. There are many variations which may be better suited to implementation in the field.

We denote by $L = ((\tau^1, a^1), (\tau^2, a^2), z) \in (\{0,1\} \times A)^2 \times [0,1]$ a lottery over two assignments (τ^1, a^1) and (τ^2, a^2), where assignment (τ^1, a^1) is chosen with probability z. We denote by $\mathcal{L}$ the uniform distribution over such lotteries and consider i.i.d. random variables $(L_k)_{k\in\{1,2\}}$ distributed according to $\mathcal{L}$.

Consider the mechanism $\widetilde{G} = (M, \mu)$ such that $M = \mathcal{U}_0$. Given a message $m \in \mathcal{U}_0$, assignment μ is decided as follows:

1. two lotteries L_1 and L_2 are drawn according to $\mathcal{L}$
2. the experimenter picks the lottery L^* that maximizes the agent's stated preferences m
3. the agent's assignment is drawn according to lottery L^*.

Clearly, it is in the agent's interest to state his true preferences. Furthermore, every type has positive probability of being assigned to every pair $(\tau, a) \in \{0,1\} \times A$. It follows that $\widetilde{G}$ is a most informative mechanism.

A.2 General mechanisms

Throughout the paper we consider mechanisms such that the assignment of agent i depends only on the message sent by agent i. This facilitates the implementation of such mechanisms to the extent that transfers between the agents and the principal only need to occur once, and the assignments of multiple agents need not occur at the same time. This being said, although there may be significant practical hurdles, there is no theoretical difficulty in allowing the assignment of agent i to be dependent on: the outcome y_i of agent i, the outcomes y_{-i} of other agents, the messages m_{-i} sent by other agents. While considering such mechanisms is beyond the scope of our analysis, we outline here what sort of information they could elicit and the situations in which that information could be valuable.

Assignment conditional on own outcome: By allowing the agent to choose between contracts that depend on his outcome y_i, it is possible to elicit the beliefs of agent i over his outcome y_i. A side effect is that such contracts will change the agents' incentives to exert effort. This may not be a problem if contracts related to y_i simply ensure that the agent puts high effort trying to adopt the technology. However, such contracts may also lead the agent to put effort manipulating his measured outcome.

Assignment conditional on other's outcomes: Offering agents contracts indexed on other's outcomes would potentially allow to infer the agent's beliefs over how the treatment affects the returns of others. This might also provide information about what the agent perceives are the social returns to the treatment.

Assignment conditional on other's messages: Finally, one may elicit an agent's beliefs over others' beliefs by making his assignment dependent on others' messages. For instance on top of mechanism G^*, one could offer the agent some money to guess the average message of others in the population. This would provide useful information to understand social learning (i.e. do agents think that others put effort...).

A.3 Forward looking agents

The framework of Section 4 considers myopic agents that only maximize their current payoffs. When an agent of type t is forward looking, his value $V(t)$ for being in the treatment group simply becomes

$$V(t) = \max_{e\in E} \int_{R\in\mathcal{R}} \int_{y\in Y} v(\eta, e, y)\mathrm{d}F_y^1(R, \eta, e)\mathrm{d}\theta(R) + \delta W(t, \theta(\cdot|\mathcal{I})) \tag{10}$$

where δ is the agent's discount factor and $W(t, \theta(\cdot|\mathcal{I}))$ is the continuation value of an agent with initial type t whose beliefs $\theta(\cdot|\mathcal{I})$ have been updated after receiving information $\mathcal{I}$.

Provided that the experimental setup does not change the information $\mathcal{I}$ obtained by the agent at the end of the experiment (this will be the case if the outcome of the experiments are only privately observed), then eliciting the agent's valuation does not change the agent's effort and the analysis of Sections 5 and 6 is unchanged. Because Section 7 specifically considers the impact of information provision, the analysis there becomes more subtle.

In Section 7 we consider the effect of giving agents information about the technology. One way to provide this kind of information would be to run a public experiment. The issue is that anticipating to receive information may change the agents' incentives to put effort in the first period. Indeed, as (10) highlights, agent's incentives to put effort depend on the value of the information $\mathcal{I}$ generated by his private experimentation. The perspective of getting additional information later on may reduce the agent's incentive to experiment. Chamley and Gale (1994) study this point in detail in a dynamic model of investment.

B Proofs

Proof of Proposition 1: We break down the data $\mathbf{d}_G$ in two subsamples $(d_G^{\sigma_1(i)})_{i\in\mathbb{N}}$ and $(d_G^{\sigma_0(i)})_{i\in\mathbb{N}}$ such that σ_0, σ_1 are non-decreasing mappings from $\mathbb{N}$ to $\mathbb{N}$, and for all $i \in \mathbb{N}$, $\tau_{\sigma_1(i)} = 1$ and $\tau_{\sigma_0(i)} = 0$. Since $\forall m \in \cup_{t\in T} Rat^G(t), \mathbb{E}_\mu(\tau_i|m_i = m) \in [\xi, 1-\xi]$, we have that each such subsample is infinite and we can pick σ_1 and σ_0 to be strictly increasing from $\mathbb{N} \to \mathbb{N}$. We define mapping h such that $h(\mathbf{d}_G) \sim \mathbf{d}_{G_0}$ as follows.

We use the notation $h(\mathbf{d}_G) = (d_i^h)_{i\in\mathbb{N}}$, where $d_i^h = (m_i^h, p_i^h, \tau_i^h, y_i^h)$. For every $i \in \mathbb{N}$, we set $m_i^h = \emptyset$, $p_i^h = 0$, we draw τ_i as the Bernoulli variable of parameter π_0. Finally we set $y_i^h = y_{\sigma_{\tau_i}(i)}$. It is easy to check that indeed, $h(\mathbf{d}_G) \sim \mathbf{d}_{G_0}$. ■

Proof of Proposition 2: Since an agent's outcome is independent of the message she sends, any strategy is weakly dominant, in particular truthful revelation of her own type. ■

Proof of Proposition 3: The proof is very similar to that of Proposition 1. Consider a mechanism $G = (M, \mu_G)$ such that every player has a strictly dominant strategy. We show that $G \preceq G^*$. An agent with value $V(t_i)$ chooses a message m_i to solve

$$\max_{m \in M} \quad \mathbb{E}_\mu(\tau_i | m_i = m) V(t_i) - \mathbb{E}_\mu(p_i | m_i = m).$$

This problem is entirely defined by player i's value $V(t_i)$. Since players have strictly dominant strategies, this problem must have a unique solution. Hence we can define the mapping $h_M(V(t_i)) \equiv \arg\max_{m \in M} \mathbb{E}_\mu(\tau_i | m_i = m) V(t_i) - \mathbb{E}_\mu(p_i | m_i = m)$. By definition $h_M(V_i) \in M$. We now construct a mapping $h : \mathcal{D} \to \Delta(\mathcal{D})$ such that the data generated by G can be simulated from data generated by G^* using mapping h. For simplicity we describe the mapping h in the case where M is finite. Given $\mathbf{d}_{G^*}$, $h(\mathbf{d}_{G^*})$ is generated iteratively as follows.

First, we break down the basic data $\mathbf{d}_{G^*}$ in $2 \times \operatorname{card} M$ subsets, according to treatment τ and the message m corresponding to the value declared by the agent. Formally, for all $m \in M$ and $\tau \in \{0, 1\}$, we define $(d_{G^*}^{\sigma_{m,\tau}(i)})_{i \in \mathbb{N}}$ ordered subsequence such that for all i, $h_M(V_{\sigma_{m,\tau}(i)}) = m$ and $\tau_{\sigma_{m,\tau}(i)} = \tau$. Since $(\overline{V} - V_{\max})/\overline{V} \in (0, 1)$, all these subsamples are infinite. Hence, $\sigma_{m,\tau}$ can be chosen to be strictly increasing from $\mathbb{N} \to \mathbb{N}$. We use these subsamples to simulate data $\mathbf{d}_G$.

Let us denote $h(\mathbf{d}_{G^*}) = (d_i^h)_{i \in \mathbb{N}}$. For all $i \in \mathbb{N}$, $d_i^h = (m_i^h, p_i^h, \tau_i^h, y_i^h)$. We first set $m_i^h = h_M(V_i)$. Then using $\mu_G(m_i^h)$, we draw values τ_i^h and p_i^h. Finally we set $y_i^h = y_{\sigma_{m_i^h, \tau_i^h}(i)}$. This defines $h : \mathcal{D} \to \Delta(\mathcal{D})$. It is easy to check that $h(\mathbf{d}_{G^*}) \sim \mathbf{d}_G$.[36] This concludes the proof. ■

[36] Note that for the sake of notational simplicity, this construction ends up wasting data points by not taking consecutive elements from the subsamples. This is inconsequential here since we have infinitely many data points.

Proof of Proposition 4: Consider the probabilities $\mathbb{E}_\mu(\tau_i|m_i = m)$ for all m in M. Since every message is uniquely rationalizable for some type, it must be that for all $m \neq m'$, $\mathbb{E}_\mu(\tau_i|m_i = m) \neq \mathbb{E}_\mu(\tau_i|m_i = m')$. Otherwise, either messages m and m' are equivalent for all types, or one message dominates the other for all types. Hence, there exists a mapping $g : M \to \{1, \cdots, n\}$ such that when we define $\pi_k = \mathbb{E}_\mu(\tau_i|m_i = g^{-1}(k))$ for $k \in \{1, \cdots, n\}$ we have $\pi_1 < \cdots < \pi_n$.

In what follows we assume without loss of generality that $M = \{1, \cdots, n\}$ and $g(k) = k$ for all k. We define $p(k) = \mathbb{E}_\mu(p_i|m_i = k)$. For any $k > 1$, let us consider $V(t_i|m_i = k)$ the value of a type t_i choosing to send message k and $V(t_i|m_i = k-1)$ the value of a type choosing to send message $k-1$. We have that

$$\begin{aligned} \pi_k V(t_i|m_i = k) - p(k) &\geq \pi_{k-1} V(t_i|m_i = k) - p(k-1) \\ \pi_k V(t_i|m_i = k-1) - p(k) &\leq \pi_{k-1} V(t_i|m_i = k-1) - p(k-1). \end{aligned}$$

Subtracting the two inequalities yields that $(\pi_k - \pi_{k-1})(V(t_i|m_i = k) - V(t_i|m_i = k-1)) \geq 0$, which implies that $V(t_i|m_i = k) \geq V(t_i|m_i = k-1)$. Let us denote by V_k the maximum value such that a type of that value $V(t_i)$ would send message k. We define $V_0 = -\overline{V}$. It must be that an agent i sends message k if and only if her value $V(t_i)$ belongs to $[V_{k-1}, V_k]$. Furthermore, since $\pi_k - \pi_{k-1} > 0$, agent i has strict incentives to send message m whenever $V(t_i) \in (V_{k-1}, V_k)$.

The second part of the proposition is immediate given quasilinear preferences. ■

Proof of Proposition 5: Given mechanism $G = (M, \mu)$, we define mechanism $G' = (M, \mu')$

as follows:

$$\forall m \in M, \quad \mu'(m) = \begin{cases} \tau = 0, p = 0 & \text{with probability} \quad \underline{\rho} \\ \mu(m) & \text{with probability} \quad \overline{\rho} - \underline{\rho} \\ \tau = 1, p = 0 & \text{with probability} \quad \overline{\rho} \end{cases}$$

Clearly mechanism G' is strategically equivalent to mechanism G. The proof that $G \preceq G'$ is omitted since it is simple and essentially identical to that of Propositions 1 and 3. ■

Proof of Proposition 6: For any message $m \in M$, we define $\pi_m = \mathbb{E}_\mu(\tau_i | m_i = m)$ and $p(m) = \mathbb{E}_\mu(p_i | m_i = m)$. Consider two messages m and m' respectively sent by types with values $V(t)$ and $V(t')$. We must have that

$$\begin{aligned} \pi_m V(t) - p(m) &\geq \pi_{m'} V(t) - p(m') \\ \pi_m V(t') - p(m) &\leq \pi_{m'} V(t') - p(m'). \end{aligned}$$

These two inequalities yield that $(\pi_{m'} - \pi_m)V(t) \leq p(m') - p(m) \leq (\pi_{m'} - \pi_m)V(t')$, which implies that $|p(m') - p(m)| < (\overline{\rho} - \underline{\rho})V_{\max}$. Hence the difference in utilities between sending two messages m and m' for an agent with value $V \in [-V_{\max}, V_{\max}]$ is $|(\pi_m - \pi_{m'})V - p(m) + p(m')| \leq 2(\overline{\rho} - \underline{\rho})V_{\max}$. ■

Proof of Proposition 8: The proof of Proposition 8 is immediate and hence omitted. ■

Proof of Proposition 9: Since $G' \preceq G$, there must exist $h : \mathcal{D} \to \Delta(\mathcal{D})$ such that

$h(\mathbf{d}_G) \sim \mathbf{d}_{G'}$. We have,

$$\begin{aligned}
\mathbb{E}_P[\Delta_i(\mathbf{d}_G)] - \mathbb{E}_P[\Delta_i(\mathbf{d}_{G'})] &= \mathbb{E}_P[\Delta_i(\mathbf{d}_G)) - \Delta_i(h(\mathbf{d}_G))] \\
&= \mathbb{E}_P[\mathbb{E}_P[\theta_i(\mathcal{R}_H|\mathbf{d}_G) - \theta_i(\mathcal{R}_H|h(\mathbf{d}_G))|h(\mathbf{d}_G)]].
\end{aligned}$$

We focus on term $\mathbb{E}_P[\Delta_{i,h}(\mathbf{d}_G)|h(\mathbf{d}_G)]$ where $\Delta_{i,h} \equiv \theta_i(\mathcal{R}_H|\mathbf{d}_G) - \theta_i(\mathcal{R}_H|h(\mathbf{d}_G))$. We have

$$\Delta_{i,h} = \theta_i^h(\mathcal{R}_H)\theta_i^h(\mathcal{R}_L)\frac{f_i(\mathbf{d}_G|\mathcal{R}_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|\mathcal{R}_L, h(\mathbf{d}_G))}{\theta_i^h(\mathcal{R}_H)f_i(\mathbf{d}_G|\mathcal{R}_H, h(\mathbf{d}_G)) + \theta_i^h(\mathcal{R}_L)f_i(\mathbf{d}_G|\mathcal{R}_L, h(\mathbf{d}_G))} \tag{11}$$

where $\theta_i^h \equiv \theta_i(\cdot|h(\mathbf{d}_G))$. Furthermore,

$$\int_{\mathbf{d}_G} \Delta_{i,h} \times \left[\theta_i^h(\mathcal{R}_H)f_i(\mathbf{d}_G|\mathcal{R}_H, h(\mathbf{d}_G)) + \theta_i^h(\mathcal{R}_L)f_i(\mathbf{d}_G|\mathcal{R}_L, h(\mathbf{d}_G))\right] \mathrm{d}F(\mathbf{d}_G) = 0. \tag{12}$$

Subtracting (12) from (11) and integrating out yields that

$$\mathbb{E}_P[\Delta_{i,h}|h(\mathbf{d}_G)] = \int_{\mathbf{d}_G} \Delta_{i,h} \times \alpha_i(\mathbf{d}_G)\mathrm{d}F(\mathbf{d}_G) \tag{13}$$

where

$$\alpha_i(\mathbf{d}_G) \equiv f_P(\mathbf{d}_G|\mathcal{R}_H, t_i, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|\mathcal{R}_H, h(\mathbf{d}_G)) + \theta_i^h(\mathcal{R}_L)[f_i(\mathbf{d}_G|\mathcal{R}_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|\mathcal{R}_L, h(\mathbf{d}_G))].$$

It is straightforward to verify that whenever the principal has weakly optimistic conditional beliefs, then $\Delta_{i,h} \times \alpha_i(\mathbf{d}_G) \geq 0$ and when the principal has strongly pessimistic conditional beliefs, then $\Delta_{i,h} \times \alpha_i(\mathbf{d}_G) \leq 0$. This, in combination with (13), concludes the proof. ■

Proof of Proposition 10: Since $G' \preceq G$, there must exist $h : \mathcal{D} \to \Delta(\mathcal{D})$ such that

$h(\mathbf{d}_G) \sim \mathbf{d}_{G'}$. This implies that $W(G') = \mathbb{E}_P[\phi(\theta_i(\mathbf{d}_{G'}), t_i)] = \mathbb{E}_P[\phi(\theta_i(h(\mathbf{d}_G)), t_i)]$. Hence,

$$\begin{aligned} W(G) - W(G') &= \mathbb{E}_P[\phi(\theta_i(\mathbf{d}_G), t_i) - \phi(\theta_i(h(\mathbf{d}_G)), t_i)] \\ &= \mathbb{E}_P[\mathbb{E}_P[\phi(\theta_i(\mathbf{d}_G), t_i) - \phi(\theta_i(h(\mathbf{d}_G)), t_i)|h(\mathbf{d}_G)]]. \end{aligned} \tag{14}$$

where the last line follows from the law of iterated expectations. Focus on the term $\Delta W \equiv \mathbb{E}_P[\phi(\theta_i(\mathbf{d}_G), t_i) - \phi(\theta_i(h(\mathbf{d}_G)), t_i)|h(\mathbf{d}_G)]$. As $\theta_i(\mathbf{d}_G) \in [\underline{\theta}, \overline{\theta}]$, it follows that $\theta_i(h(\mathbf{d}_G)) \in [\underline{\theta}, \overline{\theta}]$. We denote $\Delta_{i,h}(\mathbf{d}_G) \equiv \theta_i(\mathbf{d}_G) - \theta_i(h(\mathbf{d}_G))$. Using a Taylor series expansion with Lagrange remnant, we know that there exists $\tilde{\theta} \in [\theta_i(h(\mathbf{d}_G)), \theta_i(\mathbf{d}_G)]$ such that

$$\begin{aligned} \phi(\theta_i(\mathbf{d}_G), t_i) - \phi(\theta_i(h(\mathbf{d}_G)), t_i) &= \frac{\partial \phi(\cdot, t_i)}{\partial \theta}(\theta_i(h(\mathbf{d}_G)))\Delta_{i,h}(\mathbf{d}_G) + \frac{1}{2}\frac{\partial^2 \phi(\cdot, t_i)}{\partial \theta^2}(\tilde{\theta})\Delta_{i,h}(\mathbf{d}_G)^2 \\ &\geq \frac{\partial \phi(\cdot, t_i)}{\partial \theta}(\theta_i(h(\mathbf{d}_G)))\Delta_{i,h}(\mathbf{d}_G) + \frac{1}{2}\Delta_{i,h}(\mathbf{d}_G)^2 \inf_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial^2 \phi(\cdot, t_i)}{\partial \theta^2} \end{aligned}$$

where the last line follows from the definition of the infimum and the fact that as $\theta_i(\mathbf{d}_G) \in [\underline{\theta}, \overline{\theta}]$, it follows that $\theta_i(h(\mathbf{d}_G)) \in [\underline{\theta}, \overline{\theta}]$. For concision, we denote $a = \frac{\partial \phi(\cdot, t_i)}{\partial \theta}(\theta_i(h(\mathbf{d}_G)))$ and $b = \inf_{\theta \in [\underline{\theta}, \overline{\theta}]} \frac{\partial^2 \phi(\cdot, t_i)}{\partial \theta^2}$. Thus

$$\Delta W \geq \mathbb{E}_P\left[a\Delta_{i,h}(\mathbf{d}_G) \; + \frac{b}{2}\Delta_{i,h}(\mathbf{d}_G)^2 \middle| h(\mathbf{d}_G) \right].$$

Furthermore, we have that $\theta_i(\mathbf{d}_G) = \mathbb{E}_i(\mathbf{1}_{R=R_H}|\mathbf{d}_G) = \mathbb{E}_i(\mathbf{1}_{R=R_H}|\mathbf{d}_G, h(\mathbf{d}_G))$. Hence, using the notation $\theta_i^h = \theta_i(h(\mathbf{d}_G))$, we obtain

$$\Delta_{i,h}(\mathbf{d}_G) = \theta_i^h(1-\theta_i^h)\frac{f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) + (1-\theta_i^h) f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))}.$$

Recalling that F_i and F_P are absolutely continuous with respect to F, we have

$$\begin{aligned}\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)|h(\mathbf{d}_G)) &= \int_{\mathbf{d}_G} \theta_i^h(1-\theta_i^h)\frac{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))} \\ &\quad \times f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))\mathrm{d}F(\mathbf{d}_G) \qquad (15)\end{aligned}$$

$$\begin{aligned}\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)^2|h(\mathbf{d}_G)) &= \int_{\mathbf{d}_G} [\theta_i^h(1-\theta_i^h)]^2\left(\frac{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))}\right)^2 \\ &\quad \times f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))\mathrm{d}F(\mathbf{d}_G). \qquad (16)\end{aligned}$$

Furthermore, we have

$$\int_{\mathbf{d}_G} \Delta_{i,h}(\mathbf{d}_G)[\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))]\mathrm{d}F(\mathbf{d}_G)=0. \qquad (17)$$

Using (17) with (15), we obtain

$$\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)|h(\mathbf{d}_G)) = \int_{\mathbf{d}_G} \theta_i^h(1-\theta_i^h)\frac{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))}\alpha_i(\mathbf{d}_G)\mathrm{d}F(\mathbf{d}_G) \qquad (18)$$

where

$$\alpha_i(\mathbf{d}_G)=f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)[f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))]. \qquad (19)$$

Focus on the case where $f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))\geq 0$. Then, the assumption of weakly optimistic beliefs implies that $f_p(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))\geq f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))$, and thus

$$\begin{aligned}&\alpha_i(\mathbf{d}_G)[f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))] \\ &\quad \geq \left(\frac{f_P(\mathbf{d}_G|R_H,h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))}-\theta_i^h\right)[f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))]^2.\end{aligned}$$

We now relate this to $\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)^2|h(\mathbf{d}_G))$. We have that

$$\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))} \leq \frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))}.$$

Finally observe that since $\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))} \geq 1$ we have

$$\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))} \leq \frac{1}{\theta_i^h(1-\theta_i^h)}\left(\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))}-\theta_i^h\right).$$

Consider now the case where $f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G)) \leq 0$. The assumption of weakly optimistic beliefs implies that $f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G)) \leq f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))$. Hence

$$\alpha_i(\mathbf{d}_G)[f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))] \geq (1-\theta_i^h)[f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))-f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))]^2.$$

Furthermore, we have that

$$\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H,h(\mathbf{d}_G))+(1-\theta_i^h)f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))} \leq 1 \leq \frac{1}{\theta_i^h}.$$

Altogether, this implies that $\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)|h(\mathbf{d}_G)) \geq \mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)^2|h(\mathbf{d}_G))$. Hence whenever $a+b/2 \geq 0$ then $W(G)-W(G') \geq 0$. Given that $a \geq \inf_{\theta\in[\underline{\theta},\overline{\theta}]} \frac{\partial\phi(\cdot,t_i)}{\partial\theta} \geq 0$, this concludes the proof. ■

Proof of Proposition 11: The proof is similar to that of Proposition 10 and we use the same notation. Using (14) from Proposition 10 and a Taylor series expansion with Lagrange remnant, there exists $\tilde{\theta} \in [\theta_i(h(\mathbf{d}_G)), \theta_i(\mathbf{d}_G)]$ such that

$$\begin{aligned}
\phi(\theta_i(\mathbf{d}_G),t_i) - \phi(\theta_i(h(\mathbf{d}_G)),t_i) &= \frac{\partial\phi(\cdot,t_i)}{\partial\theta}(\theta_i(h(\mathbf{d}_G)))\Delta_{i,h}(\mathbf{d}_G) + \frac{1}{2}\frac{\partial^2\phi(\cdot,t_i)}{\partial\theta^2}(\tilde{\theta})\Delta_{i,h}(\mathbf{d}_G)^2 \\
&\leq \frac{\partial\phi(\cdot,t_i)}{\partial\theta}(\theta_i(h(\mathbf{d}_G)))\Delta_{i,h}(\mathbf{d}_G) + \frac{1}{2}\Delta_{i,h}(\mathbf{d}_G)^2 \sup_{\theta\in[\underline{\theta},\overline{\theta}]} \frac{\partial^2\phi(\cdot,t_i)}{\partial\theta^2}
\end{aligned}$$

For concision, we denote $a = \frac{\partial \phi(\cdot,t_i)}{\partial \theta}(\theta_i(h(\mathbf{d}_G)))$ and $b = \sup_{\theta \in [\underline{\theta},\overline{\theta}]} \frac{\partial^2 \phi(\cdot,t_i)}{\partial \theta^2}$, thus

$$\Delta W \leq \mathbb{E}_P \left[a\Delta_{i,h}(\mathbf{d}_G) \; + \frac{b}{2}\Delta_{i,h}(\mathbf{d}_G)^2 \middle| h(\mathbf{d}_G) \right].$$

Using (16), (18) and (19) from Proposition 10, we first focus on the case $f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G)) \leq 0$. The assumption of strongly pessimistic beliefs implies that $f_p(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G)) \geq f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))$ thus,

$$\begin{aligned}
&\alpha_i(\mathbf{d}_G)[f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))] \\
&\qquad \leq -\left(\frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))} - (1-\theta_i^h)\right)[f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))]^2.
\end{aligned}$$

We now relate this to $\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)^2|h(\mathbf{d}_G))$. We have that

$$\frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) + (1-\theta_i^h) f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))} \leq \frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{(1-\theta_i^h) f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))}.$$

Finally observe that since $\frac{f_P(\mathbf{d}_G|R_H,t_i,h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_L,h(\mathbf{d}_G))} \geq 1$ we have

$$\frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{(1-\theta_i^h) f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))} \leq \frac{1}{\theta_i^h(1-\theta_i^h)}\left(\frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))} - (1-\theta_i^h)\right).$$

Let us now turn to the case where $f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G)) \geq 0$. The assumption of strongly pessimistic beliefs implies that $f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G)) \leq f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))$. Hence

$$\alpha_i(\mathbf{d}_G)[f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))] \leq -\theta_i^h[f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) - f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))]^2.$$

Furthermore, we have that

$$\frac{f_P(\mathbf{d}_G|R_H, t_i, h(\mathbf{d}_G))}{\theta_i^h f_i(\mathbf{d}_G|R_H, h(\mathbf{d}_G)) + (1-\theta_i^h) f_i(\mathbf{d}_G|R_L, h(\mathbf{d}_G))} \leq 1 \leq \frac{1}{1-\theta_i^h}.$$

Altogether, this implies that $-\mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)|h(\mathbf{d}_G)) \geq \mathbb{E}_P(\Delta_{i,h}(\mathbf{d}_G)^2|h(\mathbf{d}_G))$. Hence whenever $-a + b/2 \leq 0$ then $W(G) - W(G') \leq 0$. Given that $a \geq \inf_{\theta \in [\underline{\theta}, \bar{\theta}]} \frac{\partial \phi(\cdot, t_i)}{\partial \theta} \geq 0$, this concludes the proof. ■

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2007): "Learning and Disagreement in an Uncertain World," Massachusetts Institute of Technology, *mimeo.*

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002): "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review*, 92(5), 1535–1558.

Ashraf, N., J. Berry, and J. M. Shapiro (2008): "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia," Harvard University, *mimeo.*

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121(2), 635–672.

Banerjee, A. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107(3), 797–817.

Banerjee, A., P. Bardhan, K. Basu, R. Kanbur, and D. Mookherjee (2005): "New Directions in Development Economics: Theory or Empirics?," *Economic and Political Weekly*, 15(40), 4328–4346.

Becker, G. M., M. H. DeGroot, and J. Marschak (1964): "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9(3), 226–232.

Bergemann, D., and S. Morris (2005): "Robust Mechanism Design," *Econometrica*, 73(6), 1771–1813.

Bergemann, D., and J. Välimäki (2005): "Monopoly Pricing of Experience Goods," Cowles Foundation Discussion Paper # 1463.

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

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100(5), 992–1026.

Bohm, P., J. Lindén, and J. Sonnegård (1997): "Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanisms vs. Markets," *The Economic Journal*, 107(443), 1079–1089.

Chamley, C., and D. Gale (1994): "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, 62(5), 1065–1085.

Cohen, J., and P. Dupas (2009): "Free Distribution or Cost-Sharing? Evidence from a Malaria Prevention Experiment in Kenya," University of California, Los Angeles *mimeo.*

Crawford, V. P., and N. Iriberri (2007): "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 75(6), 1721–1770.

Deaton, A. (2009): "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," Princeton University, *mimeo.*

Duflo, E., R. Glennerster, and M. Kremer (2008): "Using Ramdomization in Development Economics Research: A Tool Kit," in *Handbook of Development Economics, Vol. 4*, ed. by T. P. Schultz, and J. Strauss, pp. 3895–3962. Elsevier, Amsterdam.

Duflo, E., M. Kremer, and J. Robinson (2008): "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya," *American Economic Review*, 98(2), 482–488.

Dupas, P. (2009a): "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya," University of California, Los Angeles *mimeo.*

——— (2009b): "What Matters (and What Does Not) in a Households' Decision to Invest in Malaria Prevention," *American Economic Review*, forthcoming.

Dupas, P., and J. Robinson (2009): "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya," University of California, Los Angeles, *mimeo.*

Eyster, E., and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73(5), 1623–1672.

Foster, A. D., and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6), 1176–1209.

Gentzkow, M., and J. M. Shapiro (2006): “Media Bias and Reputation,” *Journal of Political Economy*, 114(2), 280–316.

Heckman, J. J. (1990): “Varieties of Selection Bias,” *The American Economic Review*, 80(2), 313–318.

Heckman, J. J., and E. Vytlacil (2005): “Structural Equations, Treatment Effects, and Econometric Policy Evaluation,” *Econometrica*, 73(3), 669–738.

Imbens, G. W., and J. D. Angrist (1994): “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62(2), 467–475.

Jadad, A. R. (1998): *Randomised Controlled Trials: A User’s Guide*. BMJ Books, London, UK.

Kamenica, E., and M. Gentzkow (2009): “Bayesian Persuasion,” University of Chicago, *mimeo*.

Karlan, D. S., and J. Zinman (2009): “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment,” *Econometrica*, *forthcoming*.

Kremer, M., and E. Miguel (2007): “The Illusion of Sustainability,” *The Quarterly Journal of Economics*, 122(3), 1007–1065.

Miguel, E., and M. Kremer (2004): “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 72(1), 159–217.

Milgrom, P., and J. Roberts (1986): “Price and Advertising Signals of Product Quality,” *The Journal of Political Economy*, 94(4), 796–821.

Moffitt, R. (2008): “Estimating Marginal Treatment Effects in Heterogeneous Populations,” Johns Hopkins University, *mimeo*.

Mullainathan, S., J. Schwartzstein, and A. Shleifer (2008): “Coarse Thinking and Persuasion,” *Quarterly Journal of Economics*, 123(2), 577–619.

Ortmann, A., and R. Hertwig (2002): “The Costs of Deception: Evidence from Psychology,” *Experimental Economics*, 5(2), 111–131.

Oster, E., and R. Thornton (2008): “Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-Up,” University of Chicago, *mimeo*.

Oster, S. M. (1995): *Strategic Management for Nonprofit Organizations: Theory and Cases*. Oxford University Press, Oxford, UK.

Piketty, T. (1995): “Social Mobility and Redistributive Politics,” *The Quarterly Journal of Economics*, 110(3), 551–584.

Rayo, L., and I. Segal (2008): "Optimal Information Disclosure," Stanford University, *mimeo.*

Roy, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

Sandroni, A., and F. Squintani (2007): "Overconfidence, Insurance, and Paternalism," *American Economic Review*, 97(5), 1994–2004.

Sethi, R., and M. Yildiz (2009): "Public Disagreement," MIT, *mimeo.*

Smith, L., and P. Sørensen (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68(2), 371–398.

Stolberg, H. O., G. Norman, and I. Trop (2004): "Randomized Controlled Trials," *American Journal of Roentgenology*, 183(6), 1539–1544.

Suri, T. (2008): "Selection and Comparative Advantage in Technology Adoption," MIT, *mimeo.*

Thornton, R. (2008): "The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment," *American Economic Review*, 98(5), 1829–1863.

Van den Steen, E. (2004): "Rational Overoptimism (and Other Biases)," *American Economic Review*, 94(4), 1141–1151.

——— (2005): "On the Origin of Shared Beliefs (and Corporate Culture)," MIT, *mimeo.*

Wilson, R. (1987): "Game Theoretic Analyses of Trading Processes," in *Advances in Economic Theory, 5th World Congress*, ed. by T. Bewley, pp. 33–70. Cambridge University Press, Cambridge, MA.

Yildiz, M. (2003): "Bargaining Without a Common Prior—An Immediate Agreement Theorem," *Econometrica*, 71(3), 793–811.