Incomplete Information and Strategic Learning in Technology Adoption and Diffusion

Jinhua Zhao[1]

August 8, 2007

[1]Associate Professor, Department of Economics, Iowa State University, Ames, IA 50011. Voice: 515-294-5857. Fax: 515-294-0221. Email: jzhao@iastate.edu. I thank Subir Bose for many helpful discussions. I also thank Glenn Ellison, Wally Huffman, and seminar participants at Iowa State, Minnesota, Arizona, Maryland, the University of Washington, and UC Davis for helpful comments. The usual disclaimer applies.

Abstract

We study an incomplete information dynamic game of a finite number of agents adopting a new technology with pure information externalities. Each agent has private information about his idiosyncratic adoption cost, but shares the same public information released by early adopters about the technology's efficiency. We show that a unique symmetric pure strategy Perfect Bayesian equilibrium exists. The adoption process may start late and stop temporarily, but ends in finite periods almost surely with a positive probability that some adopters are worse off. Strategic delay prolongs the adoption process, reduces the long-run adoption rate, and in itself can cause the diffusion path to be logistic. More efficient communication about the technology slows down early adoption, but speeds up the diffusion in later periods. More efficient communication about agent types tends to speed up early adoption.

Key Words: technology adoption and diffusion, learning, strategic delay, information externalities, real options.

1 Introduction

The adoption and diffusion of new technologies have long been studied in economics. Since the seminal works of Griliches (1957) and Mansfield (1961), economists have emphasized the rational behavior of individual firms (i.e., profit maximization), and argued that adoption occurs if the new technology proves to be more profitable than the old ones. The diffusion process of a new technology is determined by the distribution of firm characteristics that affect the profitability, the time pattern of the costs and benefits of the new technologies (due possibly to learning by using or output price changes, see for example Jovanovic and Lach (1989)), and exogenous information variations and individual learning (Jensen (1982)). Feder, Just and Zilberman (1985) and Sunding and Zilberman (2001) review this literature.

There is currently growing interest in the literature on incorporating information exchange among agents into the individual rationality models of the adoption and diffusion research. The empirical literature started with agricultural technologies (Case (1992), Foster and Rosenzweig (1995), and Besley and Case (1993, 1997)) and has expanded to medical drugs (Berndt, Pindyck and Azoulay (1999)) and computers (Goolsbee and Klenow (1999)). Relying mostly on micro-level data, these studies consistently find that rational profit-maximizing agents do respond to information released by other adopters. Further, using a structural estimation model, Besley and Case (1997) found that agents also anticipate and actively respond to future information from other adopters: They tend to strategically delay adoption to wait for other adopters' information. They found that a model with the forward-looking behavior performs better than one with the agents passively responding to existing information.

The theoretical research on the role of information exchange has mostly focused on passive responses to evolving information. For example, Feder and O'Mara (1981), Jensen (1982) and McFadden and Train (1996) study the role of information externalities when agents passively respond to the *existing* information released by early adopters. Ellison and Fudenberg (1993, 1995) consider agents with bounded rationality whose (nonstrategic) learning and decision rules are exogenously given. Kapur (1995) and Persons and Warther (1997) study strategic delay in complete information frameworks. There are a finite number of agents in Kapur (1995) so that the equilibrium exists only in mixed strategies. Persons and Warther (1997) assumed infinite number of agents, which leads to the existence of pure strategies.

Studying the strategic behavior is important for understanding the role of information exchange on adoption and diffusion. When agents strategically choose to wait for others to adopt first, being able to learn from each other may hinder, rather than help, the adoption process. Thus to assess the overall effects of learning, we need to compare the cost of strategic delay with the benefit of being able to learn from existing adopters. The literature is far from reaching any conclusion on this overall effect. For the most part, it shows how the *existing* information from other adopters helps an agent's adoption decision, but has not identified the degree to which the strategic behavior has delayed the adoption process.

In this paper, we investigate the effects of information exchange and strategic delay on technology adoption and diffusion. We consider a pure information externality problem of a finite number of agents without any network effects, that is, the agents interact only through information exchange. Information about the new technology is public: all agents share the same prior and the additional information released by early adopters. An agent's only private information is his cost of adoption (i.e., his type). When the agent adopts, he (partially) reveals his type, and releases a random signal about the technology. Unlike the information cascade literature (Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992) Choi (1997), Zhang (1997), Caplin and Leahy (1998) and Chamley and Gale (1994)), the agent does not have private signals about the technology. Thus his adoption (or nonadoption) decision in itself does not reveal any information about the technology. Information updating about both the technology and agent types is Bayesian.

There are several attractive features of our model that distinguish this paper from others in the literature. First, we are able to obtain a unique symmetric pure strategy Perfect Bayesian Equilibrium (PS-PBE) with *incomplete* information about agent types. In the context of technology adoption, pure strategies are much more attractive than mixed strategies because the latter imply that an adopter is always indifferent between adopting and waiting. Empirical estimation of such a model, especially if the structural approach is followed, has to impose a restriction that the payoff of adoption is equal to that of no-adoption. The restriction is likely to severely hurt the predictive power of any estimation model, particularly given that the agents are typically heterogeneous.

Second, the content of information exchange is also richer in our paper. Agents communicate in two *uncorrelated* dimensions: the performance of the new technology (a common value) and the agent types (private values). It is possible that agents communicate well in one dimension but not the other. For instance, they may know much about each others' types (or costs) in a closely-knitted community, but the nature of the new technology may prevent effective communication about the technology's performance.[1] This setting represents a major departure from the literature, e.g., Persons and Warther (1997). It allows us to explicitly study the effects on adoption and diffusion of learning about the new technology and learning about each other. We show that their effects are different, and can even be opposite.

Our results indicate that strategic delay itself can lead to the logistic diffusion path that is often empirically observed. As more agents adopt, the number of existing signals about the technology increases, reducing the incentive of remaining agents to wait for more signals, thereby further raising the likely number of new adopters. On the other hand, since low cost agents adopt first, as time goes on, the remaining agents are of higher costs, and given the same information, they are less likely to adopt. The interaction of the two opposing forces leads to possible logistic diffusion paths. In contrast, passive learning itself without strategic behavior can only cause the path to be concave. We also show that the adoption process can stop temporarily and can have a delayed start. It ends in finite periods almost surely and some agents may be made worse off by the availability of the new technology.

The paper is organized as follows. We specify the adoption game in Section 2, and show that a unique symmetric PS-PBE exists in Section 3. We discuss possible diffusion patterns along an average equilibrium path in Section 4. Then in Section 5, we conduct an extensive numerical study of diffusion paths for a range of distributions of agent types. We use the example to show the

[1] As an example, consider a village where farmers know each other well, so that they can easily figure out the costs of each farmer operating a new technology, say a drip irrigation system. They may, however, have a difficult time telling each other how good (or bad) the system is due to weather variations and land heterogeneity.

impacts of strategic delay, stronger prior about the new technology, and communications about the technology and about agent types. We discuss the generality and implications of the model assumptions in Section 6, and conclude the paper in Section 7.

2 A Game of Technology Adoption

Consider a community of N agents, indexed by n, identical except for their costs of adopting and handling new technologies. Currently they are using a "traditional" technology, the profit of which is normalized to zero. A new technology is introduced with a per period profit of $e > 0$, which, for simplicity, is assumed to be constant overtime. The agents have imperfect prior information about e, knowing only that it is distributed on $[e_l, e_h]$ with distribution function $F_0(\cdot)$, with $e_l > 0$. That is, the new technology has a higher efficiency than the current technology, although its payoff is uncertain.

Adopting the new technology requires sunk investment of physical and/or human capital. Different agents may face different costs, depending on the agent's ability, credit limit, the technology's fit to his needs, etc. We represent the heterogeneity by each agent's type θ, and the adoption cost of a type θ_n agent is $c_n = \theta_n c$. Thus an agent of a lower type incurs lower costs of adoption. Agent n's type is his private information. Others hold the common belief that θ_n is distributed on the interval $\Theta \equiv [\underline{\theta}, 1]$ with $\underline{\theta} > 0$, according to the distribution function $G_0(\cdot)$. The beliefs about the types of different agents are independently and identically distributed. Since the adoption cost θc is sunk and $e > 0$, the adoption decision is irreversible: once adopted, no agent is willing to abandon the new technology to switch back to the traditional one.

Agents are risk neutral, and the common discount rate is r. To rule out trivial cases of certain agent types always adopting or never adopting the new technology, we assume

Assumption 1 *(i) $\frac{e_h}{r} > c$: Even for an agent of the highest cost (with $\theta = 1$), the new technology may still be a profitable investment; (ii) $\frac{e_l}{r} < \underline{\theta}c$: Even for an agent of the lowest cost, it is still possible that the new technology is a bad investment.*

We consider the adoption decisions in discrete time, $t = 1, 2, \ldots$. At the beginning of each period, all agents who have not adopted the new technology, called the *remaining agents*, make the adoption decisions. For those who decide to adopt, say agent n, his profit e is realized at the end of that period. Agent n may or may not observe e perfectly, but others can observe e only imperfectly. In particular, they observe a "profit signal" that is imperfectly correlated with e: $p_n = q(e, \varepsilon_n)$, where $q(\cdot,\cdot)$ is continuously differentiable in both arguments and $\varepsilon_n \in [\varepsilon_l, \varepsilon_h]$ is a random variable with zero mean (thus $\varepsilon_l < 0$ and $\varepsilon_h > 0$) and a positive but finite variance. Let $\mathcal{P} \subseteq \mathcal{R}$ be the set of possible values of p_n, which is determined by e_l, e_h, ε_l, ε_h, and $q(\cdot)$. Without loss of generality, we assume $0 < \partial q/\partial e < \infty$ and $0 < \partial q/\partial \varepsilon < \infty$: the observation error distorts the truth in a certain direction, but the (marginal) distortion is bounded.

Information exchange is homogeneous across all agents. In particular, signal p_n is common to all agents other than n. That is, agent n's adoption decision carries the same amount of information to all others. Further, ε_n, $n = 1, \ldots, N$, are independently and identically distributed, with distribution function $H(\cdot)$ and density $h(\cdot)$. We thereby rule out the possibility that particular agents are closer to each other and thus receive more information from adoption decisions among them. ε is assumed to be independent of θ: how much others can learn from an adopter is independent of the adopter's type. We also assume $H(\cdot)$ and $q(\cdot)$ are such that the density function of p_n given e satisfies the monotone likelihood ratio property (MLRP).[2] That is, observing a higher p_n, other agents believe that the technology is more efficient.

The agents have common prior information about e given by $F_0(\cdot)$. If a certain agent, say n, adopts the technology, others observe p_n and update their information about e in the Bayesian fashion based on $q(\cdot)$ using $F_0(\cdot)$ and $H(\cdot)$. The posterior distribution of e then becomes the starting information in the next period. The posterior is more accurate if more agents adopt (and thus release more information about e) in this period. Therefore, early adopters provide positive information externalities to other agents.

Since adopters of the new technology use it in every period, presumably they release some information about e in all subsequent periods. For simplicity, we assume that an adopter releases

[2] Let $e_1 > e_0$. MLRP is satisfied if the ratio of density of p conditional on e, $q_\epsilon(e_1, \epsilon(p, e_1))h(\epsilon(p, e_1))/q_\epsilon(e_0, \epsilon(p, e_0))h(\epsilon(p, e_0))$, is increasing in profit signal p, where $\epsilon(p, e)$ is the inverse of $q(e, \epsilon)$ given e. Many functions of $q(\cdot)$ and $h(\cdot)$ satisfy this condition, e.g., $p = e + \epsilon$ and $h(\cdot)$ is normal.

information *only once*: Both $q(\cdot)$ and ε_n are time invariant so that signal p_n remains constant through time. Therefore, if agent n adopts in period t, others observe p_n and update their belief about e at the end of this period. Since the same p_n is observed in subsequent periods, it does not carry any additional information. This assumption is an extreme form of the more realistic situation where the information value of n's signal gradually decreases.[3] With this assumption, the incentive to delay adoption is due only to learning from future *new* adopters, rather than from past adopters.

Now we formally define our adoption game. A history is a sequence of adoption and non-adoption decisions of the agents and the realized profit signals of the adopters, e.g., in period t, a history is $I_t = (\boldsymbol{a}^\tau, \boldsymbol{p}^\tau)_{\tau=0}^{t-1}$ where $(\boldsymbol{a}^0, \boldsymbol{p}^0) = \emptyset$, $\boldsymbol{a}^\tau$ is a $n \times 1$ vector with $a_n^\tau = 1$ if agent n has adopted by period τ, and $a_n^\tau = 0$ otherwise, and $\boldsymbol{p}^\tau$ is a vector of profit signals with a dimension equal to the number of new technology users by the end of period τ. Strictly speaking, $\boldsymbol{a}^\tau$ represents the *state* of technology use: $a_n^\tau = 1$ means that in period τ, agent n is using the new technology. He first adopted the technology in period $k \leq \tau$ if $a_n^k = 1$ and $a_n^{k-1} = 0$. Let $\mathcal{I}$ be the set of the histories.

At history $I_t \in \mathcal{I}$, agents who have not adopted make adoption decisions. Let $A = \{0, 1\}$ be the action space, with 1 representing adoption and 0 otherwise. Let $s(I_t, \theta_n)$ be the probability that remaining agent n adopts at history I_t. Then the behavioral strategy of a remaining agent is $s : \mathcal{I} \times \Theta \rightarrow [0, 1]$. If some remaining agent adopts, nature then moves and picks ϵ_n according to $H(\cdot)$ for each new adopter n, such that new profit signal p_n is realized and appended to $\boldsymbol{p}^{t-1}$. Therefore, following any history, each remaining agent knows who has adopted and who has not, and observes the profit signals of those adopters.

Each history $I_t \in \mathcal{I}$ contains profit signals about e from the adopters. let $P(I_t) = \{p_n \in \mathcal{P} : a_n^{t-1}(I_t) = 1\}$ be the set of these signals. Clearly, the *order* in which the early adopters adopted the technology does not affect information about e. What matters is the unordered collection of the

[3] Realistically, the profit from using the new technology q_n may fluctuate overtime. For example, for agricultural technologies, the profit changes as weather, crops grown, and other factors change. As a result, the observed profit signal p_n should fluctuate overtime with (possibly strong) serial correlation. Then others can always learn more about e by observing the signals in each later period. But the information content of later signals decreases due to the serial correlation. In the limit, our special case arises when the serial correlation is perfect.

realized signals of these adopters. Further, once an agent adopts the new technology, he makes no further decisions and is essentially out of the adoption game: the remaining agents play the game based on the updated information about e.

For simplicity, we impose the following regularity conditions:

Assumption 2 *(i) The density functions $g_0(\cdot)$ (the prior about θ), $f_0(\cdot)$ (the prior about e) and $h(\cdot)$ (the observation error ε) are continuous and bounded above and away from zero. (ii) Bayes updating of beliefs about the types and about e applies to all information sets, even ones reached with zero probability.*

Assumption (i) is satisfied by most commonly used distribution functions. We impose (ii) following the discussion in Fudenberg and Tirole (1993) (Section 8.2.3) about multi-stage games with observed actions and incomplete information. (We also impose the other conditions they discussed.)

Let $\beta = \frac{1}{1+r}$ be the discount factor. At history I_t, the remaining agents simultaneously choose either to adopt or to wait. If agent n adopts, his expected payoff is

$$\pi(P(I_t), \theta_n) = E_{e|P(I_t)}\left(\frac{e}{r}\right) - \theta_n c, \tag{1}$$

which depends only on the realized profit signals. If he waits, his expected payoff depends on his belief about the types of the remaining agents, given by $\boldsymbol{g}_{-n}(I_t)$, and their strategies in the continuation games following I_t, $\boldsymbol{s}_{-n}$. The payoff is recursively defined as

$$v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n) = \beta E_{I_{t+1}|I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}(I_t)} \max\left\{\pi(P(I_{t+1}), \theta_n), v(I_{t+1}, \boldsymbol{g}_{-n}(I_{t+1}), \boldsymbol{s}_{-n}, \theta_n)\right\}. \tag{2}$$

A perfect Bayesian equilibrium (PBE) of this game includes the strategies $\boldsymbol{s}$ and beliefs $\boldsymbol{g}(\cdot)$ so that (i) strategies are sequentially rational: at each history $I \in \mathcal{I}$, the strategies of the remaining agents in all continuation games are best responses to each other; and (ii) the beliefs are consistent with Bayes' rule at every possible history $I \in \mathcal{I}$.

3 The Unique Symmetric PS-PBE of the Adoption Game

We consider only *symmetric* equilibria where all agents adopt the same strategy $s(I, \theta)$. We show in this section that there is a unique pure-strategy PBE (PS-PBE). First, we establish a monotonicity result: at any history, an agent's payoff of adoption relative to waiting is decreasing in his type.

Proposition 1 *At any history $I_t \in \mathcal{I}$ with beliefs $\boldsymbol{g}_{-n}(I_t)$ and strategies of other agents $\boldsymbol{s}_{-n}$, the relative adoption payoff of a remaining agent n, $\pi(P(I_t), \theta_n) - v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n)$, is continuous and decreasing in his type θ_n.*

All proofs are given in the Appendix. The benefit of not adopting now is that in the future, with more information released by new adopters, agent n can avoid "bad" investment where the expected profit of adoption is less than the sunk cost. As his type θ_n increases, his expected profit of adoption given the same information decreases linearly. Consequently, the probability of bad investment increases. The two forces working together imply that the relative benefit of adopting now decreases in θ_n.

Proposition 1 implies that generically, agents play pure strategies at each history: $s(I_t, \theta) = 1$ if $\theta \leq \eta(I_t, \boldsymbol{g}_{-n})$ and $s(I_t, \theta) = 0$ if $\theta > \eta(I_t, \boldsymbol{g}_{-n})$, where $\eta(I_t, \boldsymbol{g}_{-n})$, the solution to $\pi(P(I_t), \cdot) - v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \cdot) = 0$, is a *critical type* such that he adopts if and only if $\theta \leq \eta(I_t, \boldsymbol{g}_{-n})$. Then Proposition 1 implies that $\pi - v > 0$ for $\theta < \eta$ (so that $s = 1$) and $\pi - v < 0$ for $\theta > \eta$ (so that $s = 0$). For $\theta = \eta$, we restrict the agent to choose $s = 1$. From Assumption 1, such a critical type always exists in the interval $[\underline{\eta}, \bar{\eta}]$ where $\underline{\eta} = \frac{e_l}{rc} < \underline{\theta}$ and $\bar{\eta} = \frac{e_h}{rc} > 1$.[4] In the balance of the paper, we represent each agent's strategy by his critical type $\eta(I_t, \boldsymbol{g}_{-n})$ for all $I_t \in \mathcal{I}$ and $\boldsymbol{g}_{-n}$.

Information updating about agent types becomes extremely simple under the critical type representation. If agent m's strategy is critical type η and he does not adopt at history I_t, the posterior of his type is simply the prior conditional on $\theta_m > \eta$. If he adopted, the posterior is simply the prior conditional on $\theta_m \leq \eta$. That is, a strategy represented by a critical type generates beliefs represented by a number at immediately subsequent histories. Therefore, given $g_0(\cdot)$, we can repre-

[4]Particularly, *if* there exists a type $\theta < \underline{\eta}$, the agent adopts with probability one because his cost, θc, is lower than any possible realization of payoff, which is higher than e_l/r. The converse applies to $\bar{\eta}$.

sent the belief at any history I_t, $\boldsymbol{g}(I_t)$, by a scalar $\hat{\eta}_t$ that equals the critical type at the immediate preceeding history. It implies that the types of all remaining agents are distributed according to $g_0(\cdot)$ conditional on $\theta \in (\hat{\eta}_t, 1]$:

$$g_t(\theta) = g_0(\theta | \theta > \hat{\eta}_t) = \frac{g_0(\theta)}{1 - G_0(\hat{\eta}_t)}. \tag{3}$$

Given the above representation of the strategies and beliefs, we can write the continuation payoff in (2) as $v(I_t, \hat{\eta}_t, \boldsymbol{\eta}, \theta_n)$ where $\hat{\eta}_t$ is the (commonly held) belief and $\boldsymbol{\eta}$ describes the (symmetric) strategies, or critical types, of all remaining agents other than n in $\Gamma(I_t)$, the continuation games following I_t. We sometimes write $\boldsymbol{\eta}$ as $(\eta, \boldsymbol{\eta}_{t+})$ where η is the critical type at I_t, and $\boldsymbol{\eta}_{t+}$ are those at subsequent histories.

Next we show that a symmetric PS-PBE of the adoption game, represented by $\eta^*(I_t, \hat{\eta}_t)$ for $I_t \in \mathcal{I}$ and $\hat{\eta}_t \in [\underline{\theta}, 1]$, exists and is unique. Part of the intuition is provided by the next Lemma, which states that in a two period decision problem, an agent has more incentive to wait if in the future there will be more information about the new technology.

Lemma 1 *Consider a two period decision problem where agent n of type θ_n can adopt the new technology now, in the next period, or never. At history $I \in \mathcal{I}$ in the current period, consider two scenarios: one where there will be j new adopters other than agent n in the current period, and two where there will be m new adopters in addition to the j adopters. Then agent n's expected payoff of waiting now is higher under scenario two than under scenario one.*

Consider next a profile of critical types $\boldsymbol{\eta}^*$, which describes the symmetric strategies at each history in $\mathcal{I}$, and the beliefs generated by $\boldsymbol{\eta}^*$, $\hat{\boldsymbol{\eta}}$. By definition, $(\boldsymbol{\eta}^*, \hat{\boldsymbol{\eta}})$ are equilibrium strategies and beliefs if, for every $I_t \in \mathcal{I}$, $\eta^*(I_t) \in [\underline{\eta}, \bar{\eta}]$, and satisfies

$$w(I_t, \hat{\eta}(I_t), \boldsymbol{\eta}^*) \equiv \pi(P(I_t), \eta^*(I_t)) - v(I_t, \hat{\eta}(I_t), (\eta^*(I_t), \boldsymbol{\eta}^*_{t+}), \eta^*(I_t)) = 0. \tag{4}$$

That is, if all other agents follow a strategy given by $\eta^*(I_t)$, an agent's best response is to follow the same critical type $\eta^*(I_t)$ as well. If such $\boldsymbol{\eta}^*$ exists, we call it the equilibrium critical types.

Next we show that if an agent's type happens to be $\eta^*(I_t)$, his decision problem can be reduced to a two period problem.

Proposition 2 *Consider a non-terminal history $I_t \in \mathcal{I}$ on an equilibrium path. Suppose $\eta_t^* = \eta^*(I_t)$ exists in (4) and $\eta_t^* \in (\hat{\eta}_t, 1]$. If an agent of type η_t^* waits at I_t, then either he adopts in the next period, or he will never adopt and the game ends in period $t+1$ with none of the remaining agents adopting.*

The intuition for this Proposition is straightforward: by the next period, the agent of type η_t^* will be of the lowest type among all remaining agents. If he finds it not worthwhile to adopt at a certain history in period $t+1$ (when the new signals released in period t indicate a sufficiently low e), by Proposition 1, no other agents will adopt and the game ends with none of the remaining agents adopting.

The two period decision characterization of the critical type η_t^* has some flavor of the "one-step property" of optimal decisions in Chamley and Gale (1994), and generalizes Persons and Warther (1997) to dynamic Bayesian games. Chamley and Gale (1994) focused on mixed strategies, and any type is by definition indifferent between investing and waiting. If all agents wait (their strategies are symmetric), no new information is released and they will wait forever. In Persons and Warther (1997), agents' adoption costs are known and all other agents wait for the lowest cost agent to adopt first. If instead he waits, all others also wait.

Proposition 2 simplifies the derivation of $v(I_t, \hat{\eta}_t, \boldsymbol{\eta}^*, \eta_t^*)$. In particular, we know that if η_t^* exists and if $\eta_t^* > \hat{\eta}_t$,

$$v(I_t, \hat{\eta}_t, \boldsymbol{\eta}^*, \eta_t^*) = \beta E_{I_{t+1}|I_t, \hat{\eta}_t, \eta_t^*} \pi(P(I_{t+1}), \eta_t^*) 1(\pi(P(I_{t+1}), \eta_t^*) \geq 0), \tag{5}$$

where $1(\cdot)$ is an indicator function returning the value of one if its argument is true, and zero otherwise. Equation (5) indicates that the continuation payoff of type η_t^* depends on other agents' critical types at history I_t, but not on their critical types at later histories. This is a direct consequence of Proposition 2: since type η_t^* never waits beyond period $t+1$, other agents' strategies in later periods do not matter. We thus write $w(\cdot)$ and $v(\cdot)$ in (4) as $w(I_t, \hat{\eta}_t, \eta_t^*)$ and $v(I_t, \hat{\eta}_t, \eta_t^*, \eta_t^*)$ respectively.

Proposition 3 *There exists a unique solution $\eta_t^* \in [\underline{\eta}, \bar{\eta}]$ to $w(I_t, \hat{\eta}_t, \eta_t^*) = 0$ in (4). That is, there exists a unique symmetric PS-PBE in the adoption game. In particular, $w(I_t, \hat{\eta}_t, \eta)$ is continuous* *and decreasing in η, with $w(I_t, \hat{\eta}_t, \underline{\eta}) > 0$ and $w(I_t, \hat{\eta}_t, \bar{\eta}) < 0$.*

Note that the equilibrium strategy η_t^* can be below the prior $\hat{\eta}_t$. Then the game ends without any remaining agents adopting. (We will discuss this point in greater detail in the next section.) Otherwise, if $\eta_t^* > \hat{\eta}_t$, all agents whose types are in $(\hat{\eta}_t, \eta_t^*]$ adopt in this period, and the starting belief at histories in the next period is simply $\hat{\eta}_{t+1} = \eta_t^*$. Next we show that

Proposition 4 *Let η_t^* be the equilibrium critical type at history I_t with belief $\hat{\eta}_t$. η_t^* is continuous and increasing in $P(I_t)$ and $\hat{\eta}_t$.*

The continuity follows from the fact that $w(I_t, \hat{\eta}_t, \eta)$ is continuous and monotone in $\hat{\eta}_t$, η and $P(I_t)$. Note that an infinitesimal change in $P(I_t)$ does not affect the "amount of information" $P(I_t)$ carries, which is determined by the dimension of $P(I_t)$, $\dim(P(I_t))$, or the number of profit signals. A small change in $P(I_t)$ involves changing the value of these signals, but not the number of them. It is therefore intuitive that the profit of adoption and the payoff of delay are both continuous in $P(I_t)$. The information level, or $\dim(P(I_t))$, can change only discretely in our model. Then we would expect discrete changes in η_t^* as the information level changes.

As $P(I_t)$ increases, the profit signals point to a higher likelihood that e is of a high value. Then adopting now becomes more attractive because it is less likely that waiting will avoid unprofitable adoption. Consequently η_t^* increases. Higher $\hat{\eta}_t$ means that the remaining agents are of higher types and are less likely to adopt the technology. Waiting then becomes less valuable since fewer new signals are likely to arise in the future. Thus adopting now becomes more attractive, or η_t^* increases.

4 The Diffusion Paths of the New Technology

In this section we study characteristics of equilibrium diffusion paths of the new technology. Obviously, a path is determined by the specific history of the game that is realized. We therefore focus on features that are shared by many paths.

4.1 Temporary Stops and Delays

A unique feature of our adoption game is that the adoption process may stop temporarily, i.e., there may be periods when no new adopters arise, before the process resumes. Specifically, it is possible that many periods pass before the new technology even *starts* to be adopted. For example, suppose $\theta_n > \eta^*(\emptyset, \underline{\theta}) \equiv \eta_1^*$ for all $n = 1, \ldots, N$. Then nobody adopts in the first period because every agent, based on his prior belief, expects with a positive probability that there are some low cost types (those with $\theta < \eta_1^*$) among the group, and each wishes to wait for their profit signals. However, observing no new adopters and realizing that all types are thus above η_1^*, agents update their beliefs in period two: $\hat{\eta}_2 = \eta_1^*$. Since $\hat{\eta}_2 > \hat{\eta}_1 \equiv \underline{\theta}$, Proposition 4 implies that $\eta_2^* > \eta_1^*$: now agents are more willing to adopt given the new belief. If still no agent adopts, the updated belief in period three is $\hat{\eta}_3 = \eta_2^* > \eta_1^* = \hat{\eta}_2$, so that $\eta_3^* > \eta_2^*$ and more types are willing to adopt. This process of updating and increasing $\hat{\eta}_t$ and η_t^* continues until period τ when $\theta_n < \eta_\tau^*$ for some agent n, that is, when these types finally realize that they are the lowest types! Then the technology starts to be adopted.

By similar arguments, it is possible for the adoption process to stop for periods and then to resume. During the stopped periods, no further information is revealed about e, but information about agent types is always updated. That is, $\hat{\eta}_t$ and thus η_t^* gradually "work themselves up" until η_t^* surpasses the lowest type of the remaining agents. The following Proposition formally establishes the existence of temporary stops.

Proposition 5 *Define the end of the adoption game as history $I_\tau \in \mathcal{I}$ such that the continuation game at I_τ is the same as all continuation games following I_τ, with the same belief about types and no new adopters. Then at any history* before the end of the game, *there is a strictly positive probability in equilibrium that no agent adopts at the history. In particular,*

(a) there is a strictly positive probability that nobody adopts in period one, or the diffusion process has a delayed start, and

(b) at any non-terminal history, there is a strictly positive probability that the diffusion process temporarily stops.

Temporary stopping and delayed start characterizes the diffusion of many technologies, especially at the level of individual communities (e.g., villages). In our model, the temporary stops are due entirely to the incomplete information about agent types. The stops are temporary because learning still takes place, although only about agent types. Temporary stopping is a unique feature of our model that is not shared by the threshold models (e.g., Caswell and Zilberman (1986) and Feder and O'Mara (1981)), common value models (e.g., Chamley and Gale (1994)), or complete information models (e.g., Persons and Warther (1997)), where the diffusion stops permanently whenever there are no new adopters in any period. In both Feder and O'Mara (1981) and Persons and Warther (1997), learning takes place only about the technology. If at any history I_t there are no new adopters, no new information is released about e and future histories will remain the same as I_t. Consequently no agent will adopt at the subsequent histories, and the diffusion completely stops at I_t. In Chamley and Gale (1994), there is no private value, and an agent's type (or his private signal) directly reflects the technology's efficiency (the common value). If nobody adopts in a period, agents would judge that the technology is not worthwhile, permanently "killing off" the diffusion process.

4.2 The Ending of the Diffusion Process

Before the game ends, the belief about agent types $\hat{\eta}_t$ increases along an equilibrium path overtime. The reason is that in any period before the end of the game, the equilibrium strategy η_t^* is higher than the belief $\hat{\eta}_t$ and $\hat{\eta}_{t+1} = \eta_t^*$. The game ends either when all agents have adopted, or when $\eta_t^* \leq \hat{\eta}_t$. In the latter scenario, nobody adopts in this period since the strategy η_t^* is below the type of every remaining agent. Further, there is no information updating about agent types because $\hat{\eta}_{t+1} = \hat{\eta}_t > \eta_t^*$ (rather than $\hat{\eta}_{t+1} = \eta_t^*$). Information about e and beliefs in period $t+1$ are the same as those in period t, and nobody adopts in period $t+1$ either. Continuing this argument, we know the game ends in period t with none of the remaining agents ever adopting.

There is a simple condition to verify $\eta_t^* \leq \hat{\eta}_t$:

Proposition 6 *At history $I_t \in \mathcal{I}$, the necessary and sufficient condition for $\eta_t^* \equiv \eta^*(I_t, \hat{\eta}_t) \leq \hat{\eta}_t$ is $\pi(P(I_t), \hat{\eta}_t) \leq 0$.*

As long as $\pi(P(I_t), \hat{\eta}_t) > 0$, the lowest perceived type $\hat{\eta}_t$ can always adopt the technology with positive expected payoff; the game will not end without this type adopting, given the current information about e in $P(I_t)$. Further, information about e changes only when there are new adopters, which would occur only when this type (and possibly other types) adopts.

The condition $\pi(P(I_t), \hat{\eta}_t) \leq 0$ can occur only in one way: when the signals of the new adopters in period $t-1$ imply extremely low e values. That is, although $\pi(P(I_{t-1}), \hat{\eta}_t) > 0$, new signals released by those who adopted in period $t-1$ are so negative that the updated belief about e is sufficiently low, resulting in $\pi(P(I_t), \hat{\eta}_t) \leq 0$. (Note that since $\hat{\eta}_t = \eta^*_{t-1}$, and since the game did not end in period $t-1$, we must have $\pi(P(I_{t-1}), \hat{\eta}_t) > 0$.) Thus the "shocks" of low signals abruptly "kill off" the adoption process: $\pi(P(I_\tau), \hat{\eta}_\tau) > 0$ for $\tau < t$, and $\pi(P(I_t), \hat{\eta}_t) \leq 0$. In other words, along an equilibrium path without complete diffusion (i.e., adoption by everybody), at every non-terminal history, the technology is believed to be profitable for at least some of the low cost types among the remaining agents. In the last period, or at the terminal history, the technology is believed to be unprofitable for every type of the remaining agents. Proposition 7 formally establishes these results. (We omit the proof, which follows almost immediately from the discussion above.)

Proposition 7 *Along an equilibrium path,*
(a) the adoption game ends either when all agents have adopted, or the first time history I_t is reached, where profit signals $P(I_t)$ and belief $\hat{\eta}_t$ are such that $\pi(P(I_t), \hat{\eta}_t) \leq 0$.
(b) Before the terminal history is reached, the prior beliefs $\hat{\eta}_\tau$, for $\tau < t$, increase in τ.

In many cases, the signals from early adopters are their realized profits. Then an immediate corollary of Proposition 7(a) is that unless every agent adopts, there is a strictly positive probability that the last adopters will be worse off. Based on information in I_t about e, types $\theta \in [\eta^*(I_t, \hat{\eta}_t), \hat{\eta}_t]$ who have adopted earlier should not have done so.

Corollary 1 *Suppose the signals of adopters are their realized profits. Then unless the new technology is adopted by everybody, with positive probability some of the last adopters are made worse off by the opportunity to adopt the new technology.*

Can it ever happen that $\pi(P(I_t), \hat{\eta}_t) > 0$ for all history I_t along an equilibrium path, without complete diffusion of the technology? The answer is yes, and when this does occur, the diffusion path lasts infinite periods. In particular,

Proposition 8 *At any nonterminal history I_t along an equilibrium path, suppose $\pi(P(I_t), \hat{\eta}_t) > 0$, and let $\tilde{\eta}_t$ be such that $\pi(P(I_t), \tilde{\eta}_t) = 0$. If none of the remaining agents' types falls in the range $(\hat{\eta}_t, \tilde{\eta}_t]$, no more adoption occurs after history I_t but the game lasts infinite periods: $\hat{\eta}_\tau$ increases in τ for all $\tau > t$, and $\lim_{\tau \to \infty} \hat{\eta}_\tau = \tilde{\eta}_t$.*

Since there are only a finite number of agents, there is a strictly positive probability of the adoption process lasting infinite periods at any non-terminal history. Notice that even though the process may last infinite periods, the long-run adoption rate is fixed at the level when the last adoption occurred, say in period t. To outside observers, it seems like the adoption process has already ended in period t. But to those who have not adopted, they have not abandoned the opportunity of adoption even though the available information indicates that they should not adopt. They they are still expecting new adopters, and remain ready to respond to the new signals to be generated by these adopters. However, as time goes by, the probability of each remaining agent adopting gradually decreases, and eventually goes to zero. As shown in the proof of Proposition 8, $\lim_{\tau \to \infty}(\eta^*_\tau - \hat{\eta}_\tau) = 0$. (Later on, Proposition 10 will show that $\eta^*_\tau - \hat{\eta}_\tau$ decreases in τ.)

Propositions 7 and 8 imply that unless the new technology is adopted by everybody, an adoption process either never ends or is killed off by sufficiently low signals in the last period. In other words,

Corollary 2 *For technologies that are not adopted by everybody, the only manner in which the adoption process ends (in finite periods) is to be killed off by sufficiently low signals in the last period.*

We next calculate the expected long-run adoption rate.

4.3 Logistic Diffusion Paths

It has long been recognized that the diffusion paths of many technologies are logistic: the number of adopters gradually increases in early periods, reaches a peak level, and gradually dies down until the end of the diffusion process. Different reasons have been proposed in the literature, such as the heterogeneity of potential adopters in threshold models with passive learning (e.g., Feder and O'Mara (1981)). In this section, we argue that strategic delay could be an important factor behind the logistic paths.

Given the agent types $\{\theta_1, \ldots, \theta_N\}$, the diffusion path is completely determined by the realizations of the profit signals of the adopters. For example, if we fix a sample selection of the observation errors of all agents, that is, if we fix the value of $\varepsilon_n, \forall n$, then the diffusion path is completely determined. Since the profit signals are random, the diffusion paths for a technology are also random and can assume a variety of shapes e.g., concave, convex, logistic, etc. Below we present two arguments to explain why "on average," the observed diffusion paths should be logistic.

First, as more agents adopt, more information about e is released, reducing the incentive of remaining agents to wait and raising their incentive to adopt. This factor tends to accelerate the adoption process, leading to a convex diffusion path. Second, since low cost agents adopt first, the remaining agents are of higher and higher cost types. Everything else equal, a higher cost agent is less willing to adopt. This factor slows down adoption over time, making the diffusion path concave. The two factors, working against each other, can then generate logistic paths.

Consider first the impact of increasing number of adopters. To single out this effect, we compare two histories I^1 and I^2 containing profit signals indicating the same expected efficiency, $E_{e|P(I^1)}(e) = E_{e|P(I^2)}(e)$, but different number of adopters, $\dim(P(I^1)) < \dim(P(I^2))$. We show that, given the same beliefs about agent types, the remaining agents are more willing to adopt at I^2, the history containing more adopters.

Proposition 9 *Consider two histories I^1 and I^2 with $E_{e|P(I^1)}(e) = E_{e|P(I^2)}(e)$ and $dim(I^1) < dim(I^2)$, and the same belief $\hat{\eta}$. Then*

$$\eta^{*1} \equiv \eta^*(I^1, \hat{\eta}) < \eta^*(I^2, \hat{\eta}) \equiv \eta^{*2}, \tag{6}$$

if both η^{*1} and η^{*2} are strictly above $\hat{\eta}$.

There are two forces underlying (6). As there are fewer remaining agents at I^2, the distribution of the number of future signals shifts to the left. Further, the value of an additional signal is lower at I^2 because the prior about e is stronger. This second point illustrates that the additional value of future information, where information is measured by the number of signals, goes down as the starting information increases. This result is consistent with Hirshleifer and Riley (1992) (page 184), which shows in a different context that the value of the new signals is lower as the prior becomes stronger.

On an "average" diffusion path, the observation errors of early adopters tend to cancel each other out. In other words, *ex ante*, the expected efficiency of the new technology should remain constant along the average path, even as more profit signals are released. Then Proposition 9 implies that the average path tends to be convex, ignoring the fact that the types of remaining agents increase along the path.[5]

Of course, this fact cannot be ignored. The next Proposition shows that at any history, the incentive to adopt decreases as the belief about agent types increases.

Proposition 10 *Consider two possible beliefs at a history $I \in \mathcal{I}$, $\hat{\eta}^i$, $i = 1, 2$, with $\underline{\theta} \leq \hat{\eta}^1 < \hat{\eta}^2 \leq 1$. Let $G_t(\eta^{i*}; \hat{\eta}^i) \equiv \frac{G_0(\eta^{i*}) - G_0(\hat{\eta}_i)}{1 - G_0(\hat{\eta}_i)}$, where $\eta^{i*} \equiv \eta^*(I, \hat{\eta}^i)$, $i = 1, 2$, be the equilibrium probabilities of adoption by each remaining agent. Then*

$$G_t(\eta^{2*}; \hat{\eta}^2) < G_t(\eta^{1*}; \hat{\eta}^1), \tag{7}$$

which implies

$$G_0(\eta^{2*}) - G_0(\hat{\eta}^2) < G_0(\eta^{1*}) - G_0(\hat{\eta}^1).$$

The intuition of the Proposition is that the remaining agents are more reluctant to adopt and more willing to wait when their types are higher. Thus to induce an agent of type η^{2*} (which is higher than η^{1*}) to be indifferent between waiting and adopting, it must be that he will expect, according to his own belief, fewer future signals (relative to an agent of type η^{1*}), which requires $G_t(\eta^{2*}; \hat{\eta}^2)$, the probability of each remaining agent adopting in this period given $\hat{\eta}^2$, to be smaller than $G_t(\eta^{1*}; \hat{\eta}^1)$.

Proposition 10 implies that as the diffusion progresses, if we ignore the fact that more informa-

[5] We need to emphasize that this reasoning is appropriate only for the average diffusion path; on individual paths, the observation errors may not cancel out at all.

tion about e is being released, the path tends to be concave as the number of new adopters in each period will go down. The shape of the diffusion path then depends on the interaction of the two forces identified in Propositions 9 and 10.

To see the interaction, consider an "average" diffusion path where *ex ante* the expected value of e remains constant as the diffusion progresses. (This would be true if the prior about e is unbiased.) Let d be the cumulative number of adopters in a period. As d increases overtime along the diffusion path, $\hat{\eta}$ rises as low cost types drop out of the game: $\hat{\eta}'(d) > 0$. The proof of Proposition 3 showed that the number of new adopters, $L(d)$, follows a Binomial distribution with $N - d$ draws and probability $G(\eta^*(d, \hat{\eta}); \hat{\eta})$, where $\eta^*(d, \hat{\eta})$ is the equilibrium strategy. Proposition 9 showed that $G(\cdot)$ is increasing in d holding $\hat{\eta}$ constant (since $G(\cdot)$ is increasing in η^*), while Proposition 10 showed that $G(\cdot)$ decreases in $\hat{\eta}$ holding d constant. The expected number of new adopters is

$$EL(d) = (N - d)G(d, \hat{\eta}(d)), \tag{8}$$

where G is rewritten directly as a function of d and $\hat{\eta}$ for easiness of presentation, with $G_d > 0$ and $G_{\hat{\eta}} < 0$.

From (8), we know

$$EL'(d) = -G(d, \hat{\eta}(d)) + (N - d)(G_d + G_{\hat{\eta}}\hat{\eta}'(d)) = G(d, \hat{\eta}(d)) \left[\left(\frac{N}{d} - 1 \right) \epsilon_d^G - 1 \right],$$

where ϵ_d^G is the elasticity of $G(\cdot)$ with respect to d. Clearly, if the factor in Proposition 10 dominates that in Proposition 9, i.e., if $G_{\hat{\eta}}\hat{\eta}'(d)$ (which is negative) dominates G_d (which is positive), $\epsilon_d^G < 0$ and the expected number of adopters in each period decreases as d rises, resulting in a concave diffusion path. But if G_d dominates $G_{\hat{\eta}}\hat{\eta}'(d)$ so that $\epsilon_d^G > 0$, then $EL'(d) > 0$ when d is low and $EL'(d) < 0$ when d is high, representing a logistic diffusion path.

Intuitively, we expect that ϵ_d^G to be positive for low values of d and negative as d becomes larger. However, it is impossible to fully and analytically characterize the shape of the path in a general model. We next turn to a numerical example to, among other objectives, illustrate how learning and strategic delay help shape the diffusion path to be logistic.

Table 1: Parameter Values

Variables	Values
1. Information about e	
mean of $F_0(\cdot)$	50
std dev of $F_0(\cdot)$	15 (10 under scenario 1: stronger priors)
true value of e	65
2. Profit signal shocks	
mean of $H(\cdot)$	0
std dev of $H(\cdot)$	20 (10 under scenario 2: more accurate signals)
3. Information about types	
true lower bound $\underline{\theta}$	20.1
true upper bound	60.1
believed lower bound of $G_0(\cdot)$	10.1 (20.1 under scenario 3: stronger belief on types)
believed upper bound of $G_0(\cdot)$	70.1 (60.1 under scenario 3: stronger belief on types)
4. Other parameters	
size of community N	50
cost coefficient c	32.42
discount rate r	0.05

5 Strategic Learning, Information, and Agent Communications

Based on the equilibrium strategies, in this section we study the effects on the equilibrium diffusion paths of a number of factors including strategic learning, prior information about efficiency e, and agent communications about their types and about e. These factors are directly related to important policies that are routinely used by governments to promote technology adoption. Our results illustrate whether these policies can in fact always serve the purposes they are designed for, and disspell some common misconceptions.

In the numerical example, we study the independent diffusion of a single technology in a number of communities. We build on the literature of technology diffusion with passive learning such as Feder and O'Mara (1981) and with complete information such as Persons and Warther (1997), and assume Gaussian learning about the technology's efficiency e. In particular, the prior $F_0(\cdot)$ is normal, the signals are additive with $p_n = e + \epsilon_n$, $n = 1, \ldots, N$, and the distribution of observation shocks $H(\cdot)$ are also normal. The major advantage of this approach is its simplicity: with conjugate priors, the posterior distribution of e is normal, with simple formula for its mean and variance.

Table 1 lists the parameter values used in the numerical example. There are two minor departures from the theoretical setup. First, since $F_0(e)$ is normal, there is a positive probability of $e < 0$. We therefore assume that adoption is irreversible even if $e < 0$, and choose the mean and variance of $F_0(\cdot)$ to minimize this probability. In our example, the mean is over three standard deviations above zero. Second, we expanded the support of the type distribution. While it is $[\underline{\theta}, 1]$ in the theory model, now we allow the upper bound to be well above 1. The sole purpose of doing so is to make the graphs easier to present and less cluttered.

One of the major arguments in threshold models for logistic diffusion paths is the existence of bell-shaped distribution of agent types.[6] To show how strategic learning can lead to logistic diffusion even without this limiting assumption about types, we allow the types to follow five kinds of Beta distributions on $[20.1, 60.1]$: $Beta(0.5, 0.5)$, $Beta(1, 1)$, $Beta(5, 5)$, $Beta(1, 2)$, and $Beta(2, 1)$. Their densities are shown in Figure 1. Clearly, other than $Beta(5, 5)$, the type densities are not bell-shaped. For these densities, logistic paths should not on average materialize according to threshold models.

For each type distribution, the corresponding belief $H(\cdot)$ follows the same distribution but on a wider support of $[10.1, 70.1]$, capturing the notion of less precise initial belief about types. Better information about types is then represented by a narrower support of $H(\cdot)$. We also allow for biased initial information about e. In our baseline setup, the mean of F_0 is one standard deviation below $e^{true} = 65$, the true efficiency of the new technology. We tried other scenarios where F_0 is unbiased or is biased upwards, and the resulting diffusion paths share similar characteristics. The adoption cost is such that the type in the middle of the support, 40.1, is indifferent between the two technologies: $c = \frac{e}{40.1r}$. When types are symmetrically distributed, *ex ante* the technology should be adopted by half of each community under perfect and complete information.

Figure 2 presents the *average* diffusion paths across 200 communities for each of the type distributions and for several variations from the baseline parameter values. The paths are calculated in the following fashion. For each community, we randomly draw the types and profit signals for the

[6] As the adoption cost goes up, the density of agents at each cost level first goes up and then goes down. The number of new adopters per period, being proportional to the density, follows a similar pattern, leading to the logistic path.

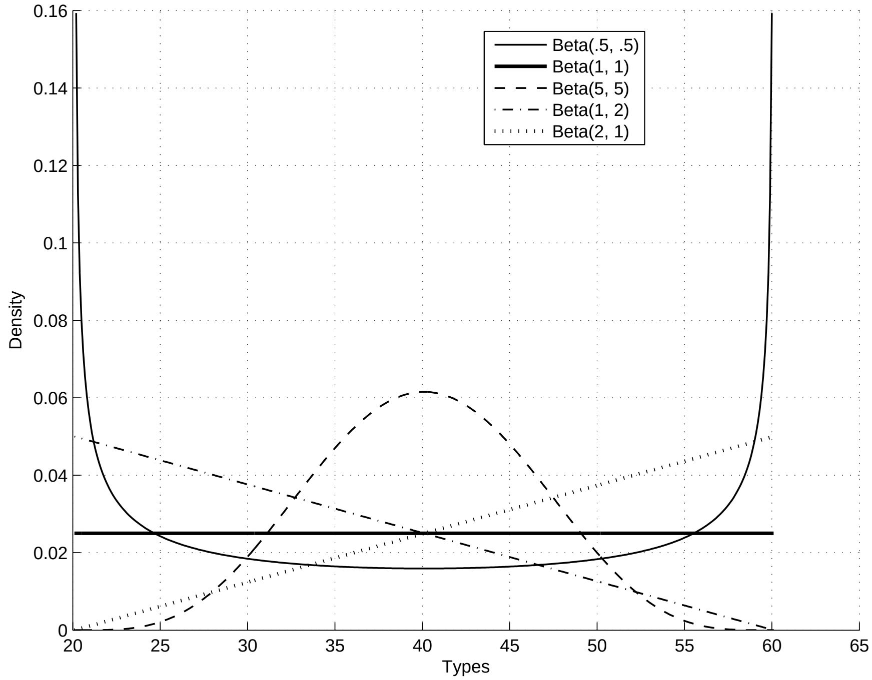


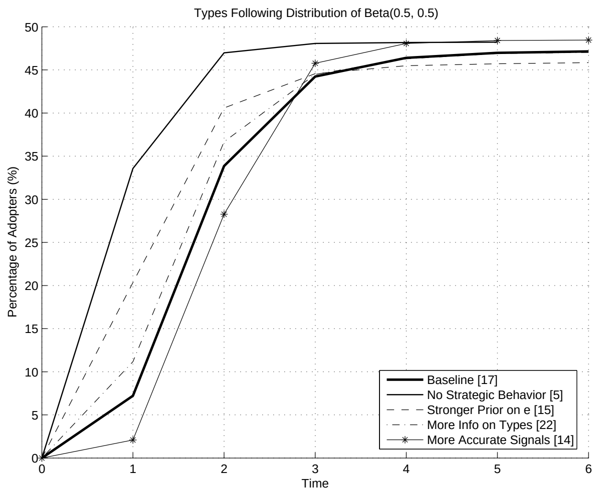
Figure 1: Densities of Type Distributions

50 agents according to the type and signal distributions. Given these types and signals, as well as initial information about e and initial beliefs about the types, we calculate the equilibrium diffusion path. We then move to the next community, where we randomly draw the types and signals again, independent of previous draws, and calculate the associated equilibrium diffusion path. Thus, although all communities face the same technology and have the same initial information and beliefs, they can be different from each other in terms of the actual agent types and profit signals. The average diffusion path then depicts the average number of adopters in each period across the 200 communities, and is represented by the solid bold line in Figure 2.[7]

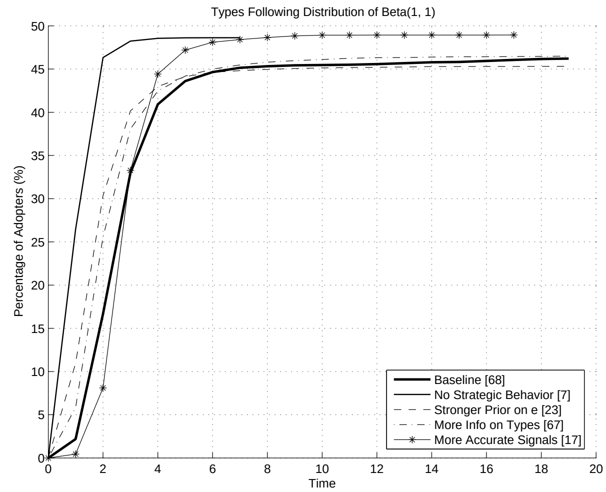
5.1 Effects of Strategic Delay

We first use the numerical results to study the effects of agents being strategic in waiting for information on the diffusion path. Thus we calculate, for each community, the diffusion path

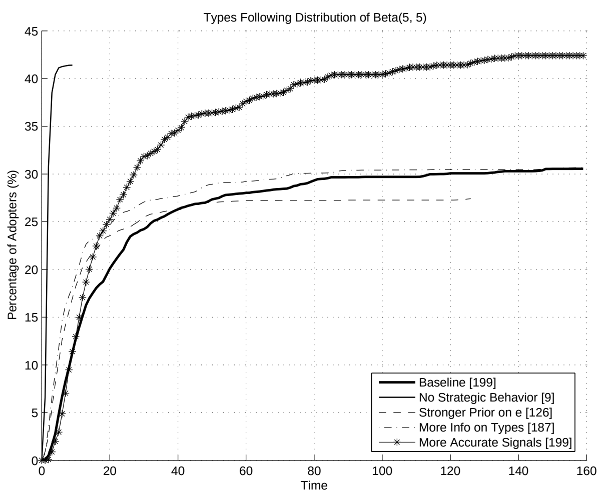
[7]The average path becomes more accurate as we increase the number of communities. However, the level of precision at 200 communities is already high: the maximum standard deviations for the average paths (maximum over all periods on each path) in the five subfigures are 0.72, 0.97, 1.67, 1.40, and 0.77 respectively. In fact, we run the program for 400 villages and the resulting average diffusion paths are nearly identical to Figure 2.



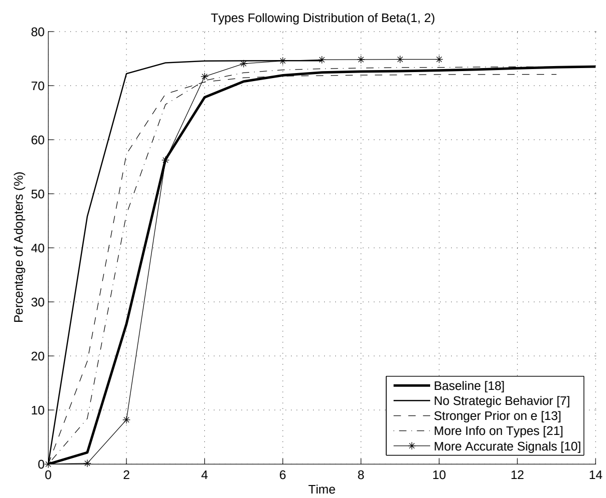
(a) U-shaped Type Densities



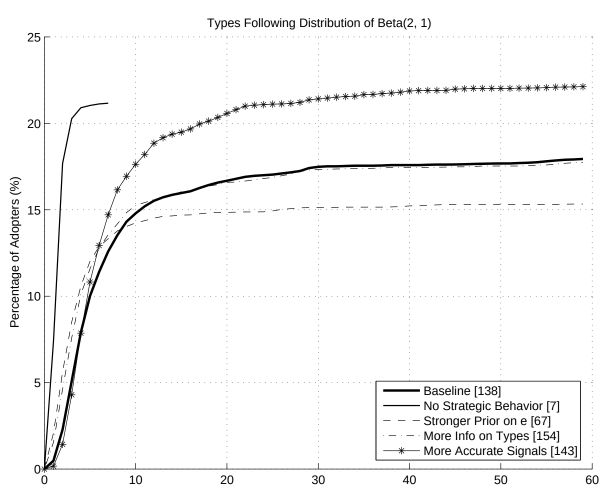
(b) Uniform Type Densities



(c) Bell-shaped Type Densities



(d) Decreasing Type Densities



(e) Increasing Type Densities

when each agent n simply makes the non-strategic adoption decision: he will adopt if and only if $E(e)/r \geq \theta_n c$, where the expectation is taken based on the current information about e. This would be the diffusion path predicted by the threshold adoption models. The non-strategic path is represented by a solid line in Figure 2. We compare these lines with the bold lined diffusion paths.

In Figure 2, the numbers in the square brackets in the legends are the periods it took for the diffusion process to stop in all 200 communities. It is then clear that for all type distributions, compared with non-strategic learning, *strategic delay prolongs the diffusion process, lowers the eventual adoption rate, and makes the diffusion path more likely to be logistic.*

The time periods it took for the diffusion process to stop with strategic learning can be more than 20 times longer than those under non-strategic learning. The difference is higher when there are heavy concentration of types who have relatively high costs and thus would not adopt in early periods, as in the case of $Beta(5,5)$ and $Beta(2,1)$. Since low cost types adopt first, there will be a large number of remaining agents left in early periods under these two type distributions, raising the expected number of signals in the future. This expectation then reduces the adoption incentive, slowing down the diffusion process. For other type distributions, e.g., $Beta(0.5,0.5)$, there is a high concentration of low cost types. Then there will be a relatively large number of adopters in early periods, raising the incentive for others to adopt both by providing more profit signals and by reducing the expected number of future signals.

For the same reason, strategic delay reduces the eventual adoption rate compared with non-strategic learning, and by a large amount when the left tails of type densities are thin as in $Beta(5,5)$ and $Beta(2,1)$. For the other type distributions, the difference between eventual adoption rates is relatively minor, implying that strategic delay mainly prolongs the diffusion process, without much reduction in the long-run adoption rate. A closer examination of the diffusion paths of individual communities showed that the prolonged delay is mainly attributable to long but temporary stops in some communities.

Without strategic delay, the diffusion paths are concave, i.e., the per period number of new adopters goes down along the diffusion path. This is true even when there are only relatively few low cost types, such as in $Beta(2,1)$ and $Beta(5,5)$. Strategic behavior induces two changes to the paths: first, diffusion in early periods is slower and gradually rises, indicating that factors in Proposition 9 dominate those in Proposition 10 for early periods; second, the order of dominance is reversed for later periods, resulting in long flat tails for the paths. This kind of long tails is rather typical of diffusion paths, e.g., the adoption of hybrid seed in Figure 1 of Griliches (1957). The two changes together cause the diffusion path to be logistic.

This kind of decline in incentive to delay as the number of signals rises can be best seen in Figure 2(a). Without any signals, the first period adoption rate with strategic delay is about 7%, while that without is 34%, a difference of 27%. However, after 34% have adopted in both cases (at the end of period two with strategic delay and period one without), the diffusion rates are respectively 44% and 47%, a difference of only 3%. In other words, the profit signals released by the early adopters, who account for 34% of the population, contain enough information to significantly reduce the remaining agents' incentives to delay and wait for further information.

5.2 Effects of Stronger Prior about e

A stylized understanding about technology diffusion is that more initial information about the technology helps speed up adoption. In the case of agricultural technology adoption, extension services of universities and government agencies, marketing specialists, and private technology consultants all play significant roles in providing prior information (Sunding and Zilberman (2001)). Although a stronger prior reduces the incentive to delay in our model, it may not always promote long term adoption rates due largely to the randomness of profit signals.

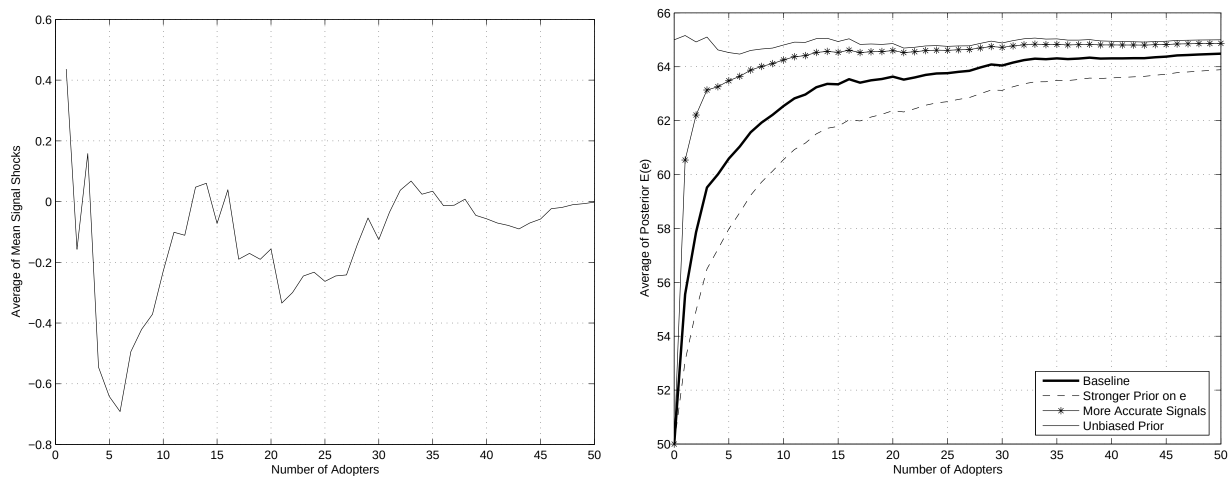
The dashed lines in Figure 2 represent diffusion paths when the standard deviation of $F_0(\cdot)$ is reduced from the baseline value of 15 to 10 (scenario 1). Since the mean of $F_0(\cdot)$ (a value of 50) is lower than the true efficiency (a value of 65), the agents have stronger but downward biased beliefs about e. We first compare the diffusion path under this prior with the baseline, to argue that under certain circumstances stronger priors always speed up diffusion even when it is downward biased.

It is clear from Figure 2 that a stronger prior always speeds up early adoption: armed with less uncertainty in beliefs about e, each agent has less incentive to wait for more signals. This is true regardless of the fact that the prior is downward biased, since even in this case, there are always low cost agents who can still benefit from adoption. These types are more willing to adopt when the prior distribution tightens.

Although the stronger prior speeds up early adoption, it reduces the eventual adoption rate since the prior is downward biased. However, it is interesting to note that the long-run adoption rate decreases only by a small amount. The reason is that profit signal released by the early adopters, which contain information about the *true* value of e, are strong enough to (nearly completely) overcome the downward bias in the prior. Figure 3(b) shows the posterior expected e as the number of signals rises, averaged across all 200 communities. Comparing the baseline with the case of unbiased prior (where F_0 is changed so that $E(F)_0 = 65$) clearly demonstrates the downward bias in posterior $E(e)$. The bias increases as the prior becomes stronger, given the same number of signals, but decreases as the number of signals rise. Since more agents adopt in early periods as the prior becomes stronger, more signals are available to overcome the downward bias in the prior, making it possible that the posterior $E(e)$ may in fact increase in a certain period. For example, at the end of period one in Figure 2(a), about 7 agents adopted in the baseline scenario and 20 agents adopted with the stronger prior. Reading from Figure 3(b), the posterior $E(e)$ under the baseline prior with seven profit signals is about 61, while that under the stronger prior with twenty signals is over 62.

Therefore, how much the long-run adoption rate can be reduced by the stronger but biased prior depends on how many agents adopt in the long run. When there are many low cost types, as in the cases of $Beta(0.5, 0.5)$, $Beta(1, 1)$ and $Beta(1, 2)$, the relatively large number of early adopters provide enough information to basically erase all the bias in the prior. In these case, the stronger prior speeds up early adoption without sacrificing the long-run diffusion rate. When the low cost types are relatively few, as in $Beta(5, 5)$ and $Beta(2, 1)$, fewer profit signals are released to overcome the bias in the prior. Consequently the long-run adoption rate decreases by a larger amount when the prior becomes stronger.

In addition to the bias in the prior, there is another reason contributing to the decrease in the long-run adoption rate under the stronger prior, namely the randomness in the signals of adopters. Figure 3(a) depicts the mean value of the random shocks as the number of adopters increases



(a) Mean Signal Shocks

(b) Posterior $E(e)$

Figure 3: Average Shocks and Posterior Expected Efficiency as Adoption Progresses

along a path, averaged over the 200 communities. In Gaussian updating about e, the mean value of the released profit signals along a diffusion path and the number of signals (given the prior) are sufficient statistics for the posterior belief about e. Differences in early adoption rates may thus lead to differences in the mean values of shocks. For instance, in the case of $Beta(2,1)$ in Figure 2(e), the long-run adoption rates are 17% and 15% for the baseline and the stronger prior scenarios, corresponding to 8.5 and 7.5 adopters. From Figure 3(a), the associated mean shock values are -0.4 and -0.5 respectively, i.e., the remaining agents observe, on average, lower profit signals under the stronger prior scenario, contributing to the lower adoption rate. In fact, when the prior is unbiased, we found that the long-run adoption rate can also decrease as the prior tightens, due to the randomness in the signals shocks.

In summary, we know

Remark 1 *(a) A stronger prior about e always speeds up early adoption, even when it is downward biased.*
(b) The stronger prior with a downward bias leads to a lower long-run diffusion rate, but by a smaller amount when there are a larger number of low cost types, and when the profit signals are more accurate.

(c) Even when the prior is unbiased, there is a positive probability that the stronger prior reduces the long-run adoption rate.

Thus, purely from the perspective of promoting adoption, extension services can be helpful even if they are "inaccurate" in the sense that they provided biased estimates of the efficiency. This is especially true if the services reduce agents' uncertainty about the technology, if there are enough early adopters, and if the profit signals of these adopters are sufficiently informative about e to overcome the initial bias.

5.3 More Information about Agent Types

Sociologists have argued that communication between new technology users and non-users helps promote adoption and diffusion. In our framework, this kind of communication can occur on two aspects: about agent types and about the new technology. Within a more closely knitted community, agents have more information about each other, i.e., their beliefs $G_0(\cdot)$ has a smaller variance σ_θ^2. However, such a community may not have a lower σ_ε^2, since the signal shocks may be caused by factors exogenous to the community, such as weather variations, market shocks, or the nature of the technology itself.[8] In our discussion, we interpret a lower σ_θ^2 as representing a higher degree of communication among agents about each other's types, while a lower σ_ε^2 as capturing more communication about the technology through more accurate profit signals.

Suppose better communication about types is represented by a mean-preserving shrinking in the spread of $G_0(\cdot)$ to $G_0'(\cdot)$. Consider history I_t with belief $\hat{\eta}_t$ where every remaining agent follows the strategy of η. Equation (8) showed that the adoption incentive increases if under the new belief, the perceived probability of adoption by others decreases: $G(\eta;\hat{\eta}_t) > G'(\eta;\hat{\eta}_t)$. Thus, to study the effects of communication about types, we only need to find the change in $G(\eta;\hat{\eta}_t)$.

Consider the starting period where $\hat{\eta}_1 = \underline{\theta}$. If the first period's equilibrium strategy under G_0, η_1^*, is lower than $\bar{\theta}$, i.e., if the first period adoption rate is lower than 50%, the probability adoption by each agent goes down given η_1^*. Then the incentive to adopt increases: better communication

[8]It may be easier to relate the output or profit of a firm to the technology used for some type of technologies than for other types.

about types raise the initial adoption rate if the technology is "gradually" adopted. For this kind of technology, only agents with relatively low costs will adopt given the original uncertainty σ_θ^2. Lower uncertainty shifts the beliefs about the distribution of agent types away from the tails, leading to a lower assessment of the probabilities of low types. That is, each agent believes that it is less likely that other agents are of really low costs. Each expects fewer adopters in the first period and is more willing to adopt. This result is reversed if $\eta_1^* > \bar{\theta}$.

The technology in our numerical example is gradually diffused, and as Figure 2 shows, the adoption rate is higher in early periods as the belief about types becomes stronger: the dash-dotted lines are the paths when the belief $G_0(\cdot)$ is distributed on $[20.1, 60.1]$, the true type space, rather than $[10.1, 70.1]$ in the baseline. The long-run diffusion rate, however, is not much affected by the stronger belief. As we discussed earlier, the adoption process stops before complete diffusion only when the realized profit signals are low. Stops due to "incorrect" expectations about the number of of low cost types are temporary, because eventually some agents will realize they are of the lowest types and start to adopt. That is, *ex ante*, or on average, changes in belief about types can only affect the speed instead of the eventual level of diffusion. We summarize the findings in

Remark 2 *(a) For technologies that are diffused gradually, better communication about agent types speeds up early adoption.*
(b) Better communication about types does not affect the ex ante *or average long-run adoption rate.*

The findings provide evidence in support of the sociologists' call for more communication among potential adopters. However, we have to caution that the communication speeds up adoption only in early periods. If the adoption process has already progressed to a degree such that more accurate $G_0(\cdot)$ increases the perceived probability of others adopting, the communication may in fact slow down the adoption rate. Communication about agent types speeds up adoption, but only if it is done at the beginning of the diffusion process.

5.4 Communication about the Technology

A major argument for increasing the interaction between adopters and potential adopters is that the latter can learn from the former about the technology, thereby becoming more willing to adopt. In the context of our model, increased communication about the technology corresponds to a lower variance in the signal shocks, σ_ε^2. A key observation, shown in Kihlstrom (1984), is that lower σ_ε^2 implies that the profit signals of the adopters carry more information about e in the sense of Blackwell (1951, 1953). Then *given the same prior information* about e, these signals are more valuable and agents will have more incentive to wait for them. Therefore, at the beginning of the diffusion process, the adoption rate should go down as σ_ε^2 becomes smaller.

Of course, once agents start to adopt, the signals from early adopters become more informative as communication becomes more efficient. The remaining agents thus have better information about e and less incentive to wait for further information. We thus expect that the speed of adoption catches up in later periods.

Figure 2 confirms the above reasoning. The rate of adoption in the case of more accurate signals is clearly lower in early periods compared with the baseline, and more so when there are many low cost types, e.g., for $Beta(0.5, 0.5)$, $Beta(1, 1)$ and $Beta(1, 2)$. Expecting a large number of early adopters, each agent's incentive to wait for signals increases a lot more when the signals become more informative. In contrast, there are only few low cost types under $Beta(5, 5)$ and $Beta(2, 1)$. Since at the beginning, agents expect few signals anyway, their incentive to wait for them only rises a little as the signals become more informative.

Figure 2 also shows that the adoption speed catches up quickly with more accurate signals, especially when there are many early adopters as in cases of $Beta(0.5, 0.5)$, $Beta(1, 1)$ and $Beta(1, 2)$. To see the reason, notice from Figure 3(b) that in this case, the posterior expected e quickly stabilizes. In other words, information about e quickly becomes so strong that further signals do not contribute much to the variation in $E(e)$ anymore. Consequently, the remaining agents are less willing to wait for them.

Since the prior is biased downward, more accurate signals also help overcome this bias, thereby raising the long-run adoption rate. From Figure 2, the increase in long-run adoption is the highest for cases of $Beta(5,5)$ and $Beta(2,1)$. Since there are only a few early signals, their increased informativeness is most helpful in overcoming the bias in the prior. In the other three type distributions, the large numbers of early signals provide enough information about e to overcome the bias, making added informativeness less valuable. In summary, we know

Remark 3 *(a) More efficient communication about the technology slows down early adoption and speeds up later adoption, especially when there are a large number of low cost types.*
(b) More efficient communication also helps overcome any bias in prior information about e*, particularly when there are few low cost types.*

Our findings thus suggest that increased communication about the technology should be encouraged only after a certain rate of diffusion. If it occurs too early, it may encourage strategic delay and slow down adoption. This risk is higher when there are many low cost types.

6 Discussion

The technology adoption and diffusion process is complex, subject to network externalities, learning by doing, certain degrees of reversibility, leap-frogging, etc. As we discuss below, some of these factors have been shown in the literature to contribute to the logistic diffusion pattern. We only focused on one aspect of the problem: the active, or strategic, information exchange among the likely adopters. In this regard, our paper relates more closely to the information cascades literature than to adoption. We discuss in this section how our results will change (or not change) as we add other factors to our model.

The costs and benefits of adoption are typically changing overtime after the introduction of a new technology. For instance, the costs can decrease due to learning by doing and increased market size of the new technology. The benefits can decrease if the increased output drives down market prices (or the marginal value of adoption) as more people adopt. Quirmbach (1986) and Jovanovic and Lach (1989) show that the two forces themselves can lead to a logistic diffusion path. Vettas (1998) allows both learning by adopters (or consumers in his paper) about the technology (or product) and by suppliers (or sellers) of the technology about the capacity of the market. The resulting prices and adoption payoffs can also lead to the logistic path. Adding these factors to our model will strengthen the likelihood of a logistic path, but will not change the role played by strategic learning. In particular, the order of adoption, or the result that low cost agents will adopt first, remains valid. However, the two period representation of the critical type's decision may break down as future costs and benefits change even without new adopters appearing. The analysis thus will be much more complicated.

In our paper, the agents are heterogeneous only in their costs of adopting the new technology. A more realistic situation would involve agents different in their profit of using the traditional as well as the new technologies. To the extent that the adoption decision is irreversible, we can let the sunk cost θc to include the profit of using the traditional technology: abandoning this technology means that its profit is lost forever. Those who are better at using the traditional technology thus face *higher* adoption costs. This is consistent with the assumptions of the leap-frogging literature (Karp and Lee (forthcoming)). Our model remains the same, except that the classification of agent types is richer.

However, including heterogeneity in using the new technology will greatly complicate our game. Suppose the profit is γe, where γ is an agent's type (private information). Every time a new profit signal is released, the remaining agents will have to update their beliefs about both e and γ. Further, the signals of earlier adopters will have to be re-assessed, given the better information about γ. That is, whenever a new signal is generated, the remaining agents have to use the entire *ordered* collection of existing signals to update their information about e and γ. The learning process becomes nearly impossible to analyze. However, the nature of learning remains the same, and given the same beliefs about γ and e, low cost agents will adopt first. The two-period simplification of the critical type's decision remains valid. Given that some ambiguities arise even in the relatively simple learning process of our model, the more complicated learning can only lead to more ambiguities, without adding much substance to our paper.

In our model, the only way of obtaining information is waiting, and the only cost of doing so is due to discounting. Many times agents actively search for information by experimentation or partially adopting a new technology (e.g., adopting a new seed on a small parcel of one's field).

Adding these factors is likely to reduce the incentive of the agents to wait for others' information, and strategic delay is likely to decrease. But unless partial adoption or experimentation is inexpensive and releases close to full information about the new technology, the incentive to delay is unlikely to disappear, and the gist of our results still holds.

7 Conclusion

In this paper, we study the overall effects of communication and learning when agents strategically delay adoption. We find that learning and strategic delay can cause the diffusion path to be logistic. More efficient communication about the technology slows down the early adoption process, even though it may speed up diffusion in later periods. On the other hand, stronger beliefs about each other's types tend to speed up early adoption and slow down the diffusion later on. On average, they do not alter the long-run adoption rate. We also find that stronger priors about the technology, even if biased downward, can speed up adoption without much reduction of the long-run adoption rate, especially when there are a large number of low cost agents.

The root of strategic delay is that the information services of adopters are external to their payoffs. A natural way of intervention is to reward early adopters by the values of their services. However, figuring out these values is more complicated than the case without the strategic behavior, as traditionally has been done.[9] A direct benefit of early adopters' information is that the information raises the payoffs of later adopters by enabling them to make better informed decisions. In addition, the information also reduces the degree of strategic delay of later adopters: armed with better information provided by the early adopters, they are less willing to (strategically) wait for more information. Specifically, at the continuation game $(I_t, \hat{\eta}_t)$, the external value of a certain signal in I_t, say p_1, is

$$\begin{aligned}\sum_{n \in \mathcal{N}_t} \Big[& \max\{\pi(I_t, \theta_n), v(I_t, \hat{\eta}_t, \eta^*(I_t, \hat{\eta}_t), \theta_n)\} \\ & - \max\{\pi(I_t \setminus p_1, \theta_n), v(I_t \setminus p_1, \hat{\eta}_t, \eta^*(I_t \setminus p_1, \hat{\eta}_t), \theta_n)\}\Big],\end{aligned} \tag{9}$$

where $\mathcal{N}_t$ denotes the set of remaining agents. *Ex ante*, before p_1 is realized, the external value of

[9] An example of the traditional approach is Shampine (1998).

agent 1 adopting in period $t-1$ is the expectation of (9) over p_1.

Proofs

Proof of Proposition 1. Let $\Gamma(I_t)$ be the continuation games following history I_t, and let $\mathcal{I}_{t+1}(I_t) \subseteq \Gamma(I_t)$ be the set of histories in period $t+1$ following I_t. Let $\mathcal{I}^1_{t+1}(I_t) \subseteq \mathcal{I}_{t+1}(I_t)$ be the subset of histories at which agent n adopts with strictly positive probability (i.e., $s(I, \theta_n) > 0$ for $I \in \mathcal{I}^1_{t+1}(I_t)$), and $\mathcal{I}^0_{t+1}(I_t)$ as that on which agent n does not adopt. Then we know

$$v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n) = \beta E_{I_{t+1}\in\mathcal{I}^1_{t+1}(I_t)}\pi(P(I_{t+1}), \theta_n) + \beta E_{I_{t+1}\in\mathcal{I}^0_{t+1}(I_t)} v(I_{t+1}, \boldsymbol{g}_{-n}(I_{t+1}), \boldsymbol{s}_{-n}, \theta_n), \tag{10}$$

where $E_{I_{t+1}\in\mathcal{I}^k_{t+1}(I_t)}$, $k = 0, 1$, are the expectations over the sets $\mathcal{I}^k_{t+1}(I_t)$ given $\boldsymbol{s}_{-n}(I_t)$ and $\boldsymbol{g}_{-n}(I_t)$, and $\boldsymbol{g}_{-n}(I_{t+1})$ is updated according to the Bayes rule. Functions $\pi(\cdot)$ and $v(\cdot)$ are the expected payoffs conditional on the histories in $\mathcal{I}^1_{t+1}$ or $\mathcal{I}^0_{t+1}$ being realized. Note that when agent n plays mixed strategies at history $I_{t+1} \in \mathcal{I}^1_{t+1}$, the expected payoffs of adopting and waiting are the same, and are represented by $\pi(P(I^1_{t+1}), \theta_n)$.

Similarly, $\pi(\cdot)$ can be rewritten as

$$\pi(P(I_t), \theta_n) = E_{I_{t+1}\in\mathcal{I}^1_{t+1}(I_t)}\pi(P(I_{t+1}), \theta_n) + E_{I_{t+1}\in\mathcal{I}^0_{t+1}(I_t)}\pi(P(I_{t+1}), \theta_n), \tag{11}$$

because $I_t = \mathcal{I}^1_{t+1}(I_t) \cup \mathcal{I}^0_{t+1}(I_t)$ and the two subsets are disjoint. Then we know

$$\begin{aligned}\pi(P(I_t), \theta_n) - v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n) &= (1-\beta)E_{\mathcal{I}^1_{t+1}(I_t)}\pi(P(I_{t+1}), \theta_n)\\ &+ E_{\mathcal{I}^0_{t+1}(I_t)}\left[\pi_n(P(I_{t+1})) - \beta v(I_{t+1}, \boldsymbol{g}_{-n}(I_{t+1}), \boldsymbol{s}_{-n}, \theta_n)\right].\end{aligned} \tag{12}$$

We can employ similar procedures to expand the last term on the right hand side of (12). Eventually, we obtain

$$\begin{aligned}\pi(P(I_t), \theta_n) - v(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n) &= (1-\beta)E_{\mathcal{I}^1_{t+1}(I_t)}\pi(P(I_{t+1}), \theta_n)\\ &+ E_{\mathcal{I}^0_{t+1}(I_t)}\Big\{(1-\beta^2)E_{\mathcal{I}^1_{t+2}(I_{t+1})}\pi(P(I_{t+2}), \theta_n)\\ &+ E_{\mathcal{I}^0_{t+2}(I_{t+1})}\Big[(1-\beta^3)E_{\mathcal{I}^1_{t+3}(I_{t+2})}\pi(P(I_{t+3}), \theta_n)\\ &+ \quad \cdots \quad \cdots\\ &+ E_{\mathcal{I}^0_{T-1}(I_{T-2})}\Big((1-\beta^{T-t})E_{\mathcal{I}^1_T(I_{T-1})}\pi(P(I^1_T), \theta_n) + E_{\mathcal{I}^0_T(I_{T-1})}\pi(P(I_T), \theta_n)\Big)\Big]\Big\},\end{aligned} \tag{13}$$

where T denotes the ending period, which may be infinity. Note that in this "last" period, the payoff of not adopting is zero. Thus only $\pi(P(I_T^0), \theta_n)$ appears in the last part of (13).

From (13), we see that the benefit of not adopting in period t arises from the potential existence of histories in $\mathcal{I}_T^0(I_{T-1})$. Facing this information set, by definition the agent would like to wait (i.e., waiting is his dominant strategy), which means that he will never adopt. That is, $\pi(P(I_T^0), \theta_n) < 0$ for $I_T \in \mathcal{I}_T^0$. Note that if $\mathcal{I}_T^0 = \emptyset$, there is no gain of waiting and agent n simply adopts at I_t.

Now we study what happens if θ_n increases. From Bayesian udpating, we know that updating about other agents' types, $\boldsymbol{g}_{-n}(I_\tau)$, is not affected by the change in θ_n at any history I_τ. If the probabilities of reaching each history is fixed, i.e., if $\mathcal{I}_\tau^k(\cdot)$ is fixed for $\tau = t+1, \ldots, T$ and $k = 0, 1$, from the expression of $\pi(\cdot)$ in (1), we know the expected profit $\pi(P(I_\tau), \theta_n)$ decreases linearly. That is, the right hand side of (13) would decrease and is linear in θ_n. However, as the expected profit of adoption decreases, since the density functions of e and ε are continuous and bounded away from zero, the set $\mathcal{I}_T^0(I_{T-1})$ becomes (weakly) larger for any I_{T-1}, and the expected value $E_{\mathcal{I}_T^0(I_{T-1})}\pi(P(I_T), \theta_n)$ decreases even if $\pi(P(I_T), \theta_n)$ remains unchanged.. That is, $E_{\mathcal{I}_T^0(I_{T-1})}\pi(P(I_T), \theta_n)$ is decreasing in θ_n. Consequently, all no-adoption sets $\mathcal{I}_\tau^0(\cdot)$ become larger. Further, this effect would again decrease the right hand side of (13). That is, $\pi(I_t, \theta_n) - v_n(I_t, \boldsymbol{g}_{-n}(I_t), \boldsymbol{s}_{-n}, \theta_n)$ decreases in θ_n. ■

Proof of Lemma 1. There will be more sample points (or profit signals) under scenario two. Further, any signals released under scenario one will also be released under scenario two. Thus the profit signals under scenario two are statistically sufficient for those under one (in inferring information about e). Since information updating is conducted in the Bayesian framework, we know the signals under scenario two are more informative than those under one in the sense of Blackwell (1951, 1953). The Lemma then follows from Kihlstrom (1984). ■

Proof of Proposition 2. Suppose the agent of type η_t^*, say agent n, waits in period t. Since η_t^* is the equilibrium critical type at history I_t, the common belief about the types of all remaining agents in period $t+1$, generated by η_t^*, is $\hat{\eta}(I_{t+1}) = \eta_t^* > \hat{\eta}_t$ for all $I_{t+1} \in \mathcal{I}_{t+1}(I_t)$.

Suppose along the equilibrium path, agent n waits at a certain history in period $t+1$, say

$I_{t+1} \in \mathcal{I}_{t+1}(I_t)$. The equilibrium strategy at I_{t+1} must satisfy $\eta^*_{t+1} \equiv \eta^*(I_{t+1}) \leq \theta_n = \eta^*_t$. For all other agents, the belief is that their types $\theta > \hat{\eta}_{t+1} = \eta^*_t \geq \eta^*_{t+1}$. That is, nobody is believed to adopt at I_{t+1}. Then information about e in period $t+2$ following I_{t+1} is the same as at I_{t+1}: $P(I_{t+2}) = P(I_{t+1})$ for all $I_{t+2} \in \mathcal{I}_{t+2}(I_{t+1})$. Further, since $\eta^*_{t+1} < \hat{\eta}_{t+1}$, every agent knows that nobody should adopt, and the beliefs about agent types is also unchanged: $\hat{\eta}(I_{t+2}) = \hat{\eta}_{t+1}$. Thus nobody adopts in period $t+2$ following I_{t+1}. Continuing this argument, we know the game ends at history I_{t+1} with none of the remaining agents adopting. That is, if agent n does not adopt at I_{t+1}, he will never adopt. ■

Proof of Proposition 3. We show the existence and uniqueness by demonstrating that $w(I_t, \hat{\eta}_t, \eta)$ is (i) continuous in η, (ii) decreasing in η, (iii) $w(I_t, \hat{\eta}_t, \underline{\eta}) > 0$, and (iv) $w(I_t, \hat{\eta}_t, \bar{\eta}) < 0$.

Consider the decision of an agent, say n, of type η when all other agents adopt the strategy represented by the critical type η at I_t. If $\eta \leq \hat{\eta}_t$, other remaining agents are not expected to adopt since their types are believed to be higher than $\hat{\eta}_t$, and $\hat{\eta}_t \geq \eta$, their critical type. Then if n does not adopt, no agents adopt in this period and the game ends: $v(I_t, \hat{\eta}_t, \eta, \eta) = \beta \max\{\pi(P(I_t), \eta), 0\}$. Then

$$w(I_t, \hat{\eta}_t, \eta) = \pi(P(I_t), \eta) - v(I_t, \hat{\eta}_t, \eta, \eta) = \min\{(1-\beta)\pi(P(I_t), \eta), \pi(P(I_t), \eta)\}. \tag{14}$$

From (1), we know $w(I_t, \hat{\eta}_t, \eta)$ is continuous and decreasing in η for $\eta < \hat{\eta}_t$.

If $\eta > \hat{\eta}_t$, strategy η generates the starting belief in the next period as $\hat{\eta}_{t+1} = \eta$. Let $L_t \leq N - \dim(P(I_t)) - 1$ be the number of *new* adopters at I_t other than agent n. We know L_t is a random variable whose distribution depends on η and $\hat{\eta}_t$ (since the types of all remaining agents are believed to be above $\hat{\eta}_t$). In particular, the probability of each agent adopting is $\text{Prob}(\hat{\eta}_t < \theta \leq \eta) = G_0(\eta) - G_0(\hat{\eta}_t) = G_t(\eta)$ (cf. (3)). Then L_t follows Binomial distribution with $G_t(\eta)$ and $N - 1 - \dim(P(I_t))$.

For each value of L_t, profit signals in the next period $P(I_{t+1})$ contains $P(I_t)$ as well as the new profit signals generated by the L_t new adopters. Thus, from (5), we know

$$v(I_t, \hat{\eta}_t, \eta, \eta) = \beta E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t} \pi(P(I_{t+1}), \eta) 1(P(\pi(I_{t+1}), \eta) \geq 0). \tag{15}$$

Assumption 2 then indicates that $v(\cdot)$ is continuous in η.

Similar to (11) and (12), we can decompose $\pi(P(I_t),\eta)$ according to L_t and I_{t+1}. Then we know

$$\begin{aligned} w(I_t,\hat{\eta}_t,\eta) =& \pi(P(I_t),\eta) - v(I_t,\hat{\eta}_t,\eta,\eta) \\ =& E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t}(1-\beta)\pi(P(I_{t+1}),\eta)1(\pi(P(I_{t+1}),\eta)\geq 0) \\ &+ E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t}\pi(P(I_{t+1}),\eta)1(\pi(P(I_{t+1}),\eta)< 0). \end{aligned} \tag{16}$$

We follow two steps to show $w(\cdot)$ decreases in η. In step one, we study $\pi(P(I_t),\theta) - u(I_t,\hat{\eta}_t,\eta,\theta)$, where

$$u(I_t,\hat{\eta}_t,\eta,\theta) = \beta E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t}\pi(P(I_{t+1}),\theta)1(\pi(P(I_{t+1}),\theta)\geq 0). \tag{17}$$

Note that $u(\cdot)$ in (17) is *not* the value of waiting of type θ when others adopt η (this value is given by $v(I_t,\hat{\eta}_t,\eta,\theta)$), since a type $\theta \neq \eta$ does not follow a *two period* decision. Rather, it measures the value of type θ *if* he decides to follow the two step decision. We use (17) only as a way of decomposing the effects of higher η in (16) into two parts: In one, captured by (17), the type of the critical type agent is higher. In the other, critical type strategies of other agents are higher. Using (17), we know

$$\begin{aligned} \pi(P(I_t),\theta) - u(I_t,\hat{\eta}_t,\eta,\theta) =& E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t}(1-\beta)\pi(P(I_{t+1}),\theta)1(\pi(P(I_{t+1}),\theta)\geq 0) \\ &+ E_{L_t|\hat{\eta}_t,\eta} E_{I_{t+1}|I_t,L_t}\pi(P(I_{t+1}),\theta)1(\pi(P(I_{t+1}),\theta)< 0). \end{aligned} \tag{18}$$

Then it is obvious that $\pi(P(I_t),\theta) - u(I_t,\hat{\eta}_t,\eta,\theta)$ decreases in θ.

Thus to show that $w(I_t,\hat{\eta}_t,\eta)$ is decreasing in η, we only need to show that $u(I_t,\hat{\eta}_t,\eta,\theta)$ is increasing in η. Lemma 1 implies that in (17), $E_{I_{t+1}|I_t,L_t}\pi(P(I_{t+1}),\theta)1(\pi(P(I_{t+1}),\theta)\geq 0)$ increases in L_t. As η increases, each agent's probability of adoption $G_t(\eta)$ increases, and the probability distribution of L_t increases in terms of first order stochastic dominance. Thus $u(I_t,\hat{\eta}_t,\eta,\theta)$ increases.

To show that $w(I_t,\hat{\eta}_t,\underline{\eta}) > 0$, note that $\pi(P(I_t),\underline{\eta}) > 0$, $\forall I_t$, from (1). Since $\underline{\eta} < \hat{\eta}_t$, (14) implies $w(I_t,\hat{\eta}_t,\underline{\eta}) > 0$. Similarly, (1) implies that $\pi(P(I_t),\bar{\eta}) < 0$, $\forall I_t$. Since $\bar{\eta} > \hat{\eta}_t$, (16) implies $w(I_t,\hat{\eta}_t,\bar{\eta}) < 0$. ■

Proof of Proposition 4. We first show that $w(I_t,\hat{\eta}_t,\eta)$ is continuous in $\hat{\eta}_t$, η and $P(I_t)$. From the proof of Proposition 3, we know $w(\cdot)$ is continuous in η. Assumption 2 and Bayesian updating imply that at each e, the conditional density $f(e|P(I))$ is continuous in $P(I)$. Thus from (1), the expected payoff $\pi(\cdot)$ is continuous in $P(I)$. Then from (15), $v(\cdot)$ is also continuous in $P(I_t)$. Therefore, $w(\cdot)$ is continuous in $P(I_t)$.

If $\eta_t^* \leq \hat{\eta}_t$, we know from (14) that $w(\cdot)$ is independent of, and thus continuous in $\hat{\eta}_t$. If $\eta_t^* > \hat{\eta}_t$, Assumption 2 and (15) imply that $v(\cdot)$ is continuous in $\hat{\eta}_t$. Since $\pi(\cdot)$ does not depend on $\hat{\eta}_t$, we know $w(\cdot)$ is continuous in $\hat{\eta}_t$.

From the definition of η_t^* and the fact that $w(\cdot)$ is decreasing in η, we know that η_t^* increases in $P(I_t)$ and $\hat{\eta}_t$ if $w(I_t, \hat{\eta}_t, \eta)$ increases in $P(I_t)$ and $\hat{\eta}_t$. We show the latter is true.

From (14), it is obvious that when $\eta_t^* \leq \hat{\eta}_t$, $w(\cdot)$ increases in $P(I_t)$ and is independent of $\hat{\eta}_t$. When $\eta_t^* > \hat{\eta}_t$, $w(\cdot)$ is given by (16). As $P(I_t)$ increases, $P(I_{t+1})$, which contains the signals in $P(I_t)$, also increases. Thus $\pi(P(I_{t+1}), \eta)$ increases, and so does $w(\cdot)$ in (16). Then $w(\cdot)$ increases in $P(I_t)$. $\pi(P(I_t), \eta)$ is independent of $\hat{\eta}_t$. As $\hat{\eta}_t$ increases, each agent's probability of adoption $G_t(\eta) = G_0(\eta) - G_0(\hat{\eta}_t)$, decreases, or the distribution of L_t decreases in the sense of first order stochastic dominance. Thus $v(\cdot)$ decreases in (15), or $w(\cdot)$ increases.

The continuity of η_t^* in $P(I_t)$ and $\hat{\eta}_t$ follows from the fact that $w(\cdot)$ is continuous and monotone in $P(I_t)$, $\hat{\eta}_t$ and η. ■

Proof of Proposition 5. Consider a non-terminal history I_t with $0 < N_t < N$ remaining agents sharing a belief of $\hat{\eta}_t$ and adopting the equilibrium strategy given by $\eta_t^* > \hat{\eta}_t$. As shown in the proof of Proposition 3, the number of new adopters at I_t, L_t, follows Binomial distribution with $G_t(\eta_t^*) = (G_0(\eta_t^*) - G_0(\hat{\eta}_t))/(1 - G_0(\hat{\eta}_t))$ and N_t. Thus, the probability that nobody adopts at I_t is $[1 - G_t(\eta_t^*)]^{N_t}$, which is strictly positive if $\eta_t^* < 1$. But since I_t is non-terminal, it must be that the equilibrium stratey η_t^* is strictly less than one. (Otherwise, every remaining agent will adopt and the game ends at I_t.) ■

Proof of Proposition 6. *Sufficiency*: Suppose $\pi(P(I_t), \hat{\eta}_t) \leq 0$. From the definition of $w(\cdot)$ in (4), the fact that $w(\cdot)$ can be written as $w(I_t, \hat{\eta}_t, \eta_t^*)$ (the argument before Proposition 3), and the fact that $v(\cdot) \geq 0$ for all values of its argument (because the worst scenario is when the agent chooses not to adopt in the next period, which results in a payoff of zero), we know $w(I_t, \hat{\eta}_t, \hat{\eta}_t) \leq 0$.

The definition of η_t^*, on the other hand, implies that $w(I_t, \hat{\eta}_t, \eta_t^*) = 0$. Since, $w(I_t, \hat{\eta}_t, \cdot)$ is decreasing (Proposition 3), we know $\eta_t^* \leq \hat{\eta}_t$.

Necessity: Suppose $\eta_t^* \leq \hat{\eta}_t$. Since $w(I_t, \hat{\eta}_t, \cdot)$ is decreasing and $w(I_t, \hat{\eta}_t, \eta_t^*) = 0$, we know $w(I_t, \hat{\eta}_t, \hat{\eta}_t) = \pi(P(I_t), \hat{\eta}_t) - v(I_t, \hat{\eta}_t, \beta_t^*, \hat{\eta}_t) \leq 0$. Since $v(\cdot) \geq 0$, we know $\pi(P(I_t), \hat{\eta}_t) \leq 0$. ■

Proof of Proposition 8. *Step 1:* show $\eta_t^* < \tilde{\eta}_t$ and $\pi(P(I_t), eta_t^*) > 0$. Suppose instead that $\eta_t^* \geq \tilde{\eta}_t$. Then from (5), $v(I_t, \hat{\eta}_t, \eta_t^*, \eta_t^*) \geq 0$ and the inequality is strict when $\eta_t^* - \tilde{\eta}_t$ is small. (When the difference is small, there is a strictly positive probability that new signals imply strictly positive adoption payoff for an agent of type η_t^*.) Thus, by definition of η_t^*, we know $\pi(P(I_t), \eta_t^*) = v(I_t, \hat{\eta}_t, \eta_t^*, \eta_t^*) \geq 0$, and the inequality is strict when η_t^* is only slightly above $\tilde{\eta}_t$. But this violates $\pi(P(I_t), \tilde{\eta}_t) = 0$ and the fact that $\pi(P(I_t), \cdot)$ is strictly decreasing. Therefore, $\eta_t^* < \tilde{\eta}_t$, which in turn implies $\pi(P(I_t), eta_t^*) > 0$.

Step 2: repeat Step 1 for $\tau > t$. Step 1 implies that in period $\tau = t+1$, we have $\pi(P(I_t), \hat{\eta}_\tau) > 0$ with $\hat{\eta}_\tau < \tilde{\eta}_t$. Since $\pi(P(I_t), \hat{\eta}_t) > 0$, Proposition 6 implies that $\hat{\eta}_{t+1} = \eta_t^* > \hat{\eta}_t$. Repeating the same argument as in Step 1, we know $\eta_\tau^* \in (\hat{\eta}_\tau, \tilde{\eta}_t)$ and $\pi(P(I_t), \eta_\tau^*) > 0$. Repeating this process for all $\tau > t$, we know $\hat{\eta}_\tau$ increases in τ, but $\hat{\eta}_\tau < \hat{\eta}_t$ for all τ.

Step 3: show $\lim_{\tau\to\infty} \hat{\eta}_\tau = \tilde{\eta}_t$. The last observation in Step 2 implies that $\lim_{\tau\to\infty}(\eta_\tau^* - \hat{\eta}_\tau) = 0$: otherwise, the monotone sequence $\hat{\eta}_\tau, \tau > t$ would be unbounded. That is, in the limit, the probability of each remaining agent adopting, $G_\tau(\eta_\tau^*)$, goes to zero: $\lim_{\tau\to\infty}(G_0(\eta_t^*) - G_0(\hat{\eta}_t))/(1 - G_0(\hat{\eta}_t)) = 0$. Thus, $\lim_{\tau\to\infty} v(I_t, \hat{\eta}_\tau, \eta_\tau^*, \eta_\tau^*) = \beta\pi(P(I_t), \eta_\tau^*)$: the expected benefit of delay is simply the discount payoff of adoption based on the current information. Then $\lim_{\tau\to\infty} w(I_t, \hat{\eta}_\tau, \eta_\tau^*) = \lim_{\tau\to\infty}(1 - \beta)\pi(P(I_t), \eta_\tau^*)$. The definition of $w(I_t, \hat{\eta}_\tau, \eta_\tau^*) = 0$, for all τ, then implies that $\lim_{\tau\to\infty} \pi(P(I_t), \eta_\tau^*) = 0$, or $\lim_{\tau\to\infty} \hat{\eta}_\tau = \tilde{\eta}_t$. ■

Proof of Proposition 9. Note that $\pi(P(I^i), \eta^{*i}) > 0$, $i = 1, 2$. Otherwise, the agent of type η^{i*} will always want to wait, instead of being indifferent between waiting and adopting. Since $\pi(\cdot)$ is continuous in η, we know $\eta^{i*} \in$ interior of the set $\{\eta \in [\underline{\eta}, \bar{\eta}] : \pi(P(I^1), \eta) = \pi(P(I^2), \eta) \geq 0\}$, $i = 1, 2$. We next show that when $\pi(P(I^i), \eta) > 0$, $w(I^i, \hat{\eta}, \eta)$ is higher with $i = 2$ than with $i = 1$. The Proposition then follows since $w(\cdot)$ is decreasing in η.

Similar to (15), we can write the expected payoff of delay as

$$v(I^i, \hat{\eta}, \eta, \eta) = \beta E_{L|\hat{\eta},\eta}\Big[E_{I_1^i|I^i,L}\pi(P(I_1^i), \eta)1(\pi(P(I_1^i), \eta) \geq 0)\Big], \tag{19}$$

where I_1^i denotes histories in the next period. As $\dim(P(I))$ increases, the prior information about e in this period is stronger. Then there will be smaller changes in the distribution of e brought forth by the L new signals. Since $\pi(I^i, \eta) \geq 0$, the possibility of "savings" by avoiding unprofitable adoption, namely the probability of $\pi(P(I_1^i), \eta) < 0$, decreases as $\dim(P(I))$ increases. That is,

$$\begin{aligned}&\pi(P(I^i), \eta) - \beta E_{I_1^i|I^i,L}\pi(P(I_1^i), \eta)1(\pi(P(I_1^i), \eta) \geq 0)\\ &= E_{I_1^i|I^i,L}\Big((1-\beta)\pi(P(I_1^i), \eta)1(\pi(P(I_1^i), \eta) \geq 0) + \pi(P(I_1^i), \eta)1(\pi(P(I_1^i), \eta) < 0)\Big)\end{aligned}$$

increases as i goes from 1 to 2. Since $\pi(P(I^i), \eta)$ is the same for $i = 1$ and $i = 2$, we know $E_{I_1^i|I^i,L}\pi(P(I_1^i), \eta)1(\pi(P(I_1^i), \eta) \geq 0)$ must decrease.

Of course, the term in the square bracket of (19) is still increasing in L. Since $N - 1 - \dim(P(I^2)) < N - 1 - \dim(P(I^1))$, the Binomial distribution of L under I^2 is to the left of that under I^1. Thus even if the value in the square bracket is the same under I^1 and I^2, $w(\cdot)$ is still lower under I^2.

Since $\pi(P(I^i), \eta)$ is the same for $i = 1$ and $i = 2$, we know $w(I^2, \hat{\eta}, \eta) > w(I^1, \hat{\eta}, \eta)$, or $\eta^{*2} > \eta^{*1}$.

■

Proof of Proposition 10. Similar to (17), we use the following expression:

$$u(I, \hat{\eta}, \eta, \theta) = \beta E_{L|\hat{\eta},\eta} E_{I^1|I,L}\pi(P(I^1), \theta)1(\pi(P(I^1), \theta) \geq 0), \tag{20}$$

where I^1 denotes the new information sets in the next period. Suppose $G_t(\eta^{2*}; \hat{\eta}^2) \geq G_t(\eta^{1*}; \hat{\eta}^1)$. Since $G_t(\cdot; \hat{\eta})$ is increasing, there exists $\xi > 0$ such that $G_t(\hat{\eta}^2 + \xi; \hat{\eta}^2) = G_t(\eta^{1*}; \hat{\eta}^1)$. and $\eta^{2*} \geq \hat{\eta}^2 + \xi$. Then

$$\begin{aligned} w(I, \hat{\eta}^1, \eta^{1*}) &= \pi(I, \eta^{1*}) - v(I, \hat{\eta}^1, \eta^{1*}, \eta^{1*})\\ &= \pi(I, \eta^{1*}) - u(I, \hat{\eta}^2, \hat{\eta}^2 + \xi, \eta^{1*})\\ &> \pi(I, \eta^{2*}) - u(I, \hat{\eta}^2, \hat{\eta}^2 + \xi, \eta^{2*})\\ &\geq \pi(I, \eta^{2*}) - v(I, \hat{\eta}^2, \eta^{2*}, \eta^{2*})\\ &= w(I, \hat{\eta}^2, \eta^{2*}).\end{aligned} \tag{21}$$

The second equality is because the distribution of L, the number of new adopters, is preserved by changing $\hat{\eta}^1$ and η^{1*} to $\hat{\eta}^2$ and $\hat{\eta}^2 + \xi$. We change $v(\cdot)$ to $u(\cdot)$ to reflect the fact that η^{1*} is *not* the critical type anymore in the second line. The strict inequality arises since $\pi(P(I), \theta) - u(I, \hat{\eta}, \eta, \theta)$ decreases in θ (see (18) in the proof of Proposition 3), and η^{1*} is replaced by $\eta^{2*} > \eta^{1*}$. The weak inequality follows since $\eta^{2*} \geq \hat{\eta}^2 + \xi$, or the distribution of L shifts to the right under η^{2*}, which implies that $v(\cdot)$ is higher under η^{2*}. Note that we change $u(\cdot)$ back to $v(\cdot)$ since η^{2*} is now the critical type as $\hat{\eta}^2 + \xi$ changes to η^{2*}.

But (21) contradicts the fact that $w(I, \hat{\eta}^1, \eta^{1*}) = w(I, \hat{\eta}^2, \eta^{2*}) = 0$, proving the Proposition. ■

References

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

Berndt, Ernst R., Robert S. Pindyck, and Pierre Azoulay, "Network Effects and Diffusion in Pharmaceutical Markets: Antiulcer Drugs," NBER Working Paper 7024 1999.

Besley, Timothy and Anne Case, "Modeling Technology Adoption in Developing Countries," *American Economic Review*, 1993, *83*, 396–402.

___ and ___, "Diffusion as a Learning Process: Evidence From HYV Cotton," Working Paper, Princeton University 1997.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, 1992, *100(5)*, 992–1026.

Blackwell, D., "The Comparison of Experiments," in "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability," University of California Press, 1951, pp. 93–102.

___, "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 1953, *24*, 265–272.

Caplin, Andrew and John Leahy, "Miracle on Sixth Avenue: Information Externalities and Search," *Economic Journal*, 1998, *108*, 60–74.

Case, Anne, "Neighborhood Influence and Technology Change," *Regional Science and Urban Economcis*, 1992, *22*, 491–508.

Caswell, Margriet and David Zilberman, "The Effects of Well Depth and Land Quality on the Choice of Irrigation Technology," *American Journal of Agricultural Economics*, 1986, *68*, 798–811.

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

Choi, Jay Pil, "Herd Behavior, the "Penguin Effect," and the Suppression of Informational Diffusion: an Analysis of Informational Externalities and Payoff Interdependency," *RAND Journal of Economics*, 1997, *28(3)*, 407–425.

Ellison, Glenn and Drew Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy*, 1993, *101*, 612–643.

___ **and** ___ , "Word-of-Mouth Communication and Social Learning," *Quarterly Journal of Economics*, 1995, *110*, 93–125.

Feder, G., R. Just, and D. Zilberman, "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, 1985, *28*, 119–133.

Feder, Gershon and Gerald T. O'Mara, "Farm Size and the Adoption of Green Revolution Technology," *Economic Development and Cultural Change*, 1981, *30*, 59–76.

Foster, Andrew and Mark Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technological Change in Agriculture," *Journal of Political Economy*, 1995, *103*, 1176–1209.

Fudenberg, Drew and Jean Tirole, *Game Theory*, Cambridge, MA: MIT Press, 1993.

Goolsbee, Austan and Peter J. Klenow, "Evidence on Learning nad Network Externalities in the Diffusion of Home Computers," NBER Working Paper 7329 1999.

Griliches, Zvi, "Hybrid Corn: an Exploration in the Economics of Technological Change," *Econometrica*, 1957, *25*, 501–522.

Hirshleifer, Jack and John G. Riley, *The Analytics of Uncertainty and Information*, Cambridge, UK: Cambridge University Press, 1992.

Jensen, Richard, "Adoption and Diffusion of an Innovation of Uncertain Profitability," *Journal of Economic Theory*, 1982, *27*, 182–193.

Jovanovic, Boyan and Saul Lach, "Entry, Exit, and Diffusion with Learning by Doing," *American Economic Review*, 1989, *79*, 690–699.

Kapur, Sandeep, "Technological Diffusion with Social Learning," *Journal of Industrial Economics*, 1995, *XLIII*, 173–195.

Karp, Larry and In Ho Lee, "Learning-by-Doing and the Choice of Technology: the Role of Patience," *Journal of Economic Theory*, forthcoming.

Kihlstrom, Richard E., "A "Bayesian" Exposition of Blackwell's Theorem on the Comparison of Experiments," in Marcel Boyer and Richard Kihlstrom, eds., *Bayesian Models in Economic Theory*, North-Holland, 1984, chapter 2, pp. 13–31.

Mansfield, E., "Technical Change and the Rate of Immitation," *Econometrica*, 1961, *29*, 741–766.

McFadden, Daniel L. and Kenneth E. Train, "Consumer's Evaluation of New Products: Learning from Self and Others," *Journal of Political Economy*, 1996, *104*, 683–703.

Persons, John C. and Vincent A. Warther, "Boom and Bust Patterns in the Adoption of Financial Innovations," *Review of Financial Studies*, 1997, *10* (4), 939–967.

Quirmbach, Herman C., "The Diffusion of New Technology and the Market for an Innovation," *Rand Journal of Economics*, 1986, *17*, 33–47.

Shampine, Allan, "Compensating for Information Externalities in Technology Diffusion Models," *American Journal of Agricultural Economics*, 1998, *80*, 337–346.

Sunding, David and David Zilberman, "The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector," in Bruce L. Gardner and Gordon C. Rausser, eds., *Handbook of Agricultural Economics, Vol 1A: Agricultural Production*, Amsterdam: Elsevier, 2001.

Vettas, Nikolaos, "Demand and Supply in New Markets: Diffusion with Bilateral Learning," *Rand Journal of Economics*, 1998, *29*, 215–233.

Zhang, Jianbo, "Strategic Delay and the Onset of Investment Cascades," *RAND Journal of Economics*, 1997, *28(1)*, 188–205.