Information Externalities and Strategic Delay
in Technology Adoption and Diffusion

Jinhua Zhao

April 13, 2001

1Assistant 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, and the University of Washington for helpful comments. The usual disclaimer applies.
Abstract

We study an infinite period 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 share the same public information, released by early adopters, about the technology’s efficiency, which is common to all agents. We show that a unique symmetric pure strategy Perfect Bayesian equilibrium exists. The adoption process may start late, stop temporarily, and last infinite periods. The diffusion path of the technology can also be logistic. More efficient communication about the technology slows down the early adoption process, but speeds up the diffusion in later periods. It may raise the long-run adoption rate of a relatively good technology, but reduce that of a mediocre technology. On the other hand, stronger beliefs about agents’ costs tend to speed up early adoption and slow down the diffusion later on. They do not alter the long-run adoption rate of extremely good technologies, but can either raise or reduce the adoption rate for other kinds of technologies.

**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 new technologies 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) reviews this literature.

Recently economists have started to incorporate 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 significant neighbor influences. That is, 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, 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. An exception is Kapur (1995), which studies strategic delay in a complete information dynamic game of identical agents. In the symmetric mixed-strategy Markov Perfect Equilibrium, agents randomize the timing
of adoption and can end up adopting at different times. Kapur (1995) finds that the diffusion path can be convex, especially in early periods.

Understanding 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. It has mostly identified 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 fact, we will show that the answer depends on the particular stage of the diffusion process.)

In this paper, we explicitly investigate the effects of information exchange and strategic delay on technology adoption and diffusion. We consider a pure information externality problem 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), Bikchandani, 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. Firstly, 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 method is used, 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.

Related to the first feature, we allow the agents to be heterogeneous in their payoffs of using the new technology. This heterogeneity, together with the strategic delay, determines the order of adoption in an equilibrium. That is, the “leaders” and “followers” much discussed in the sociology literature are due to economic forces, as argued by Griliches (1957). They do not arise by chance, as implied by a mixed strategy equilibrium (usually with identical agents and complete information).

Thirdly, 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. This setting allows us to explicitly study the effects on adoption and diffusion of learning about the new technology and learning about each other. For example, stronger beliefs about other agents’ types can delay the start of adoption and can cause temporary stops in the diffusion process. This occurs when agents guess (based on their prior information) that there are low cost agents among them and thus decide to wait. In these periods, no new information is released about the new technology, but agents can still update their beliefs about others’ types when they find nobody adopts. Learning about each other will not cause a good technology never to be adopted, because eventually some agents will realize they are the low cost types and they have to adopt first. However, consistently “bad” signals about a good technology (due to imperfect learning about the technology) can reduce the eventual adoption rate. The diffusion process stops when those who have not adopted receive sufficiently “bad” signals so that nobody adopts and everyone agrees nobody should adopt (thus no new information can be learned about the agent types).

Finally, the possibilities of delayed start and particularly temporary stops of the adoption

---

1As 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.
process implies that it may take infinite periods for the diffusion process to complete. The adoption game therefore has to be modeled as an infinite period game. This feature is different from the typical herd behavior models, because in our paper, each agent only has private information about his idiosyncratic shock (his type), but not on the common value component of the adoption payoff. For example, in Chamley and Gale (1994), each agent receives a private signal about the technology (a common value). If in any period there are not enough adopters, others would infer that the technology is not sufficiently efficient, so that nobody adopts and the game ends. Thus, there must be new adopters in each period, and with finite number of agents, the diffusion process has to stop in finite time. Doyle (2000) extends Chamley and Gale (1994) to heterogeneous agents (and thus pure strategies) and obtain similar results. Although finite period models are simpler to analyze, we believe adoption processes allowing temporary stops describe the real world better.

We develop a method that greatly simplifies the search for the PS-PBE. We show that for any type that is indifferent between adopting and waiting, his decision problem can be represented by a two period problem. Further, his optimal decision relates directly to the PS-PBE. The resulting simple representation of PS-PBE enables us to study possible diffusion paths along an equilibrium path, and to study the effects of different parameters of our problem on the diffusion pattern. Further, it provides a simple way of finding the PS-PBE numerically.

Our results indicate that information exchange itself can lead to the logistic diffusion path that is often observed empirically. 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. This factor tends to make the diffusion path convex. On the other hand, since low cost agents adopt first, as time goes on, the remaining agents are high cost ones, and given the same information, they are less likely to adopt or more willing to wait. This factor leads to a concave diffusion path. The interaction of the two opposing forces leads to possible logistic diffusion paths.

We find that, under plausible conditions, stronger beliefs about agent types is likely to speed up the early adoption process but slows down the process in later periods. In contrast, more efficient communication about the technology slows down initial adoption but speeds up diffusion in later
periods. We further identify the effects of prior information, profitability of the new technology, and the size of the community.

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 equilibrium path in Section 4. Section 5 shows how the path responds to variations in parameter values of the adoption game. 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 abilities of handling new technologies. Currently they are using a “traditional” technology, the profit of which is normalized to zero. A new technology is introduced that requires a certain sunk investment of physical and human capital. The investment cost depends on an agent’s type, denoted by $\theta$: $c_n = \theta_n c$. Thus an agent of a lower type incurs lower costs in adopting and using the technology. An agent’s type can be related to his ability, credit limit, the technology’s fit to his needs, etc. Agent $n$’s type is his private information. Others hold the common belief that $\theta_n$ 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.

The per period profit of using the new technology, which we call the technology’s efficiency, is denoted by $e$. For simplicity, we assume that $e$ is constant overtime. That is, there is no learning by doing or systemic intertemporal demand or supply shifts. (We discuss the implications of these assumptions in Section 6.) The agents have imperfect prior information about $e$, knowing only that it is distributed on $[e_l, e_h]$ with distribution function $F_0(\cdot)$. We assume $e_l > 0$. Since the adoption cost $\theta e$ is sunk, 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 the trivial cases of certain agent types always adopting or never adopting the new technology, we assume
Assumption 1 (i) $\frac{\theta}{r} > c$: Even for an agent of the highest cost (with $\theta = 1$), the new technology may still be a profitable investment; (ii) $\frac{\theta}{r} < \theta_c$: 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, $t = 0, 1, \ldots$. 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" $p_n = q(e, \varepsilon_n)$, 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.² Let $P_n \subset \mathcal{R}$, $n = 1, \ldots, N$, be the set of possible values of $p_n$, which is determined by $\varepsilon_l, \varepsilon_h, \varepsilon_n$, 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, $\varepsilon_n, n = 1, \ldots, N$, are independently and identically distributed, with distribution function $H(\cdot)$. We thereby rule out the possibility that particular agents are closer to each other and thus receive more information from adoption decisions among them. $\varepsilon$ is assumed to be independent of $\theta$: how much others can learn from an adopter is independent of the adopter’s type.

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.

²Although we use the term “signal,” our game is not a signaling game. The adopters do not strategically disclose (or withhold) their profit signals. Instead, the observation function $q(\cdot)$ is exogenously given.
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 information only once: Both \(q(\cdot)\) and \(e_n\) 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.\(^3\)

This assumption, together with Bayesian updating, implies that for those who have not adopted, the order in which the early adopters moved does not matter in terms of the information about \(e\). What matters is the unordered collection of the realized signals of these adopters.\(^4\) Thus the history of the game, or the information set, can be represented by \(I_t = \{p_n \in P_n : \text{agent } n \text{ has adopted before period } t\}\), the set of profit signals at the beginning of period \(t\). The dimension of \(I_t\) indicates how many agents have adopted prior to \(t\). Since the adoption is irreversible, we know \(I_1 \subseteq I_2 \subseteq \cdots \subseteq I_\infty\), where \(I_1 = \emptyset\) contains no signal and \(I_\infty = \{p_n \in P_n, n = 1, \ldots, N\}\) contains all possible signals.\(^5\) No new adoption occurs in period \(t\) if \(I_t = I_{t+1}\), and if some agents adopt, we must have \(I_t \subset I_{t+1}\) with \(I_{t+1} = I_t \cup \{p_n \in P_n : \text{n adopts in period } t\}\). Note that 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 \(e\).

Now we formally define our adoption game. The players at any information set \(I_t\) are agents who have not adopted prior to \(I_t\). Let \(A = \{0, 1\}\) be the action space for all agents and all periods, where action \(a = 0\) indicates not adopting and \(a = 1\) represents adopting the new technology. Let \(\mathcal{I}\) be the \(\sigma\)-algebra of the subsets of \(I_\infty\). Each element of \(\mathcal{I}\) is a possible information set the agents face.\(^6\) Let \(s(I, \theta_n)\) be the probability that agent \(n\) adopts at information \(I\). Then the behavioral

---

\(^3\)Realistically, the profit from using the new technology \(q_n\) may fluctuate over time. 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 over time 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.

\(^4\)This point will become clear after we present the Bayes formula in (2).

\(^5\)We use \(I_\infty\) because the game may last infinite periods. Of course, it can also end in finite time, say in period \(T\), in which case \(I_\infty\) should be \(I_T\). For simplicity, we use \(I_\infty\) regardless of when the game ends.

\(^6\)Here we are not particularly strict in defining \(\mathcal{I}\). We defined the information set \(I\) in terms of the realized value
strategy of an agent is \( s : \mathcal{I} \times \Theta \rightarrow [0, 1] \). Further, let \( \mathcal{I}(I) \) be the \( \sigma \)-algebra generated by \( I \), which describes all possible information sets in the continuation games of state \( I \).\(^7\)

Beliefs about other agents’ types are updated based on the observed actions according to Bayes’ rule. Suppose at the information set \( I_t \), density function \( g_m^t(\theta_m) \) describes the prior belief others hold about agent \( m \)’s type \( \theta_m \). After observing his action \( a_m \), this belief is updated as

\[
g_m^{t+1}(\theta_m | a_m, I_t) = \frac{g_m^t(\theta_m) + (1 - s(I_t, \theta_m))(1 - a_m)g_m^t(\theta_m)}{\int g_m^t(\theta_m) + (1 - s(I_t, \theta_m))(1 - a_m)dG_m^t(\theta_m)}. \tag{1}
\]

We let \( g(t) \) denote the vector of beliefs about the types of the remaining agents.

Similarly, given any prior information about \( e \), the conditional density of observing signals \( p_1, \ldots, p_n \) (which are independent) is given by \( \prod_{i=1}^{n} h(\epsilon(p_i, e)) \), where \( \epsilon(p_i, e) \) is implicitly defined by \( p_i \equiv q(e, \epsilon(p_i, e)) \). Thus the posterior density of \( e \) conditional on observing \( p_1, \ldots, p_n \) is

\[
f(e | p_1, \ldots, p_n) = \frac{\prod_{i=1}^{n} h(\epsilon(p_i, e)) f_0(e)}{\int \prod_{i=1}^{n} h(\epsilon(p_i, e))dF_0(e) \prod_{i=1}^{n} h(\epsilon(p_i, e))dF_0(e)} \tag{2}
\]

Note that the order in which the signals \( p_1, \ldots, p_n \) are revealed does not matter. For example, we can verify that (2) is equivalent to the signals arriving sequentially: after observing \( p_1 \), the agents update their beliefs about \( e \), which are again updated after observing \( p_2 \), and so on. Thus in our game, the current information set \( I \) completely describes the current beliefs about \( e \).

For simplicity, we impose the following regularity conditions:

**Assumption 2** (i) The density functions \( g_0(\cdot) \) (the prior about \( \theta \)), \( f_0(\cdot) \) (the prior about \( e \)) and \( h(\cdot) \) (the observation error \( \epsilon \)) 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.

Let \( \beta = \frac{1}{1 + r} \) be the discount factor. In any period, say \( t \), given the information set \( I_t \), the
remaining agents simultaneously choose either to adopt or to wait. If agent \( n \) adopts, his expected payoff is

\[
\pi(I_t, \theta_n) = E_{e|I_t} \left( \frac{e}{r} \right) - \theta_n c,
\]

which is independent of the actions of other agents. If he waits, his expected payoff depends on his belief about the types of the remaining agents, given by \( g_{-n}(t) \), and their strategies in periods \( t \) and later. The payoff is recursively defined as

\[
v(I_t, g_{-n}(t), s_{-n}, \theta_n) = \beta E_{I_{t+1}|I_t, g_{-n}(t), s_{-n}(I_t)} \max \{ \pi(I_{t+1}, \theta_n), v(I_{t+1}, g_{-n}(t+1), s_{-n}, \theta_n) \}.
\]

Thus \( v(I_t, g_{-n}, s_{-n}, \theta_n) \) is the *continuation payoff* at the state \( I_t \) with nonadoption. The vector \( s_{-n} \) represents the strategies of agents other than \( n \) who have not adopted before period \( t \). In particular, these strategies represent these agents’ responses to the state \( I_t \), as well as to all possible states in the continuation games \( I(I_t) \) with the corresponding beliefs determined by the Bayes rule. We slightly abuse our notation and use \( s_{-n} \) also to represent the specific responses facing a particular information set, as in the case of \( s_{-n}(I_t) \).

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

### 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 information set, an agent’s payoff of adoption relative to waiting is decreasing in his type.

**Proposition 1** At any information set \( I_t \in \mathcal{I} \) with beliefs \( g_{-n}(t) \), and given the strategies of other agents \( s_{-n} \) in the continuation games \( I(I_t) \), the relative adoption payoff of a remaining agent \( n \),

\[
\pi(I_t, \theta_n) - v(I_t, g_{-n}(t), s_{-n}, \theta_n),
\]

is continuous, decreasing and concave in his type \( \theta_n \).
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 $\theta$ 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 and is concave in $\theta_n$.

Proposition 1 implies that the behavioral strategy $s(I, \theta)$ has a pure strategy representation: $s(I, \theta) = 1$ if $\theta \leq \eta(I)$ and $s(I, \theta) = 0$ if $\theta > \eta(I)$, where $\eta(I)$ is a critical type such that he adopts if and only if $\theta \leq \eta(I)$. Particularly, $\eta(I_t)$ is the solution to $\pi(I_t, \cdot) - v(I_t, g_{-n}(t), s_{-n}, \cdot) = 0$. Then Proposition 1 implies that $\pi - v > 0$ for $\theta < \eta(I_t)$ (so that $s = 1$) and $\pi - v < 0$ for $\theta > \eta(I_t)$ (so that $s = 0$). For $\theta = \eta(I_t)$, we restrict the agent to choose $s = 1$. From Assumption 1, such a critical type always exists in the interval $[\underline{\theta}, \overline{\theta}]$ where $\underline{\theta} = \frac{\omega_1}{\omega} < \theta$ and $\overline{\theta} = \frac{\omega_n}{\omega} > 1$.$^8$

Information updating about agent types becomes extremely simple under this pure strategy. From (1), if agent $m$ does not adopt at information set $I_t$, the posterior density of his type $g_{m+1}^t(\cdot)$ is simply the prior density $g_m^t(\cdot)$ conditional on $\theta_m > \eta(I_t)$. If he adopted, the posterior is simply the prior conditional on $\theta_m \leq \eta(I_t)$. Therefore, given $g_0(\cdot)$, we can represent the starting belief at any continuation game $I_t \in I$, $g(t)$, by a scalar $\hat{\eta}_t$, which implies that the types of all remaining agents are distributed according to $g_0(\cdot)$ conditional on $\theta \in (\hat{\eta}_t, 1)$:

$$g_t(\theta) = g_0(\theta|\theta > \hat{\eta}_t) = \frac{g_0(\theta)}{1 - G_0(\hat{\eta}_t)}. \quad (5)$$

Under the new representation, we write the continuation payoff in (4) as $v(I_t, \hat{\eta}_t, \eta, \theta_n)$ where $\hat{\eta}_t$ is the belief and $\eta$ is the (symmetric) strategies, or critical types, of remaining agents other than $n$ in this period. Their strategies in later periods are determined along the equilibrium path.

Next we show that a symmetric PS-PBE of the adoption game, represented by $\eta^*(I_t, \hat{\eta}_t)$ for $I_t \in I$ 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.

$^8$ Particularly, if there exists a type $\theta < \overline{\theta}$, the agent adopts with probability one because his cost, $\theta c$, is lower than any possible realization of payoff, which is higher than $e_1/r$. The converse applies to $\underline{\theta}$. 

10
Lemma 1 Consider a two period decision problem where agent \( n \) of type \( \theta_n \) can adopt the new technology now, in the next period, or never. Given an information set \( I \in \mathcal{I} \) 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.

The Lemma implies that in a two period game with more than two remaining agents, an agent, say \( n \), has more incentive to wait, or will choose a lower \( \eta \) if he expects more others to adopt in this period. We can extend the Lemma to show that agent \( n \)'s optimal decision \( \eta \) decreases in the strategies (or \( \eta \)'s) of other agents. In principle, we can then extend the downward slopping response function to a game of more than two periods. In this case, more agents adopting in the current period will generate more signals in all future periods. To see this, consider two scenarios where an agent adopts in period \( t_1 \) and period \( t_2 \) respectively, with \( t_1 < t_2 \). If he adopts in \( t_1 \), his profit signal will always be available in period \( t_2 \). However, this approach is complicated, especially for deriving some of the comparative statics results that we are interested in.

Instead, we rely on an observation that will greatly simplify the search for the PS-PBE. Specifically, for a particular type of agent, his optimal adoption problem is equivalent to a two period decision problem: if he waits, either he adopts in the next period or he will never adopt. Further, this agent’s optimal decision helps us pin down the equilibrium strategy of every remaining agent.

First, we use Proposition 1 to identify this particular type of agent.

Proposition 2 Given any information set \( I_t \in \mathcal{I} \) with belief \( \hat{\eta}_t \in [\underline{\eta}, \bar{\eta}] \), and the dimension of \( I_t \) \( \text{dim}(I_t) < N \), suppose a number \( \eta^*_t \in [\underline{\eta}, \bar{\eta}] \) exists, satisfying

\[
w(I_t, \hat{\eta}_t, \eta^*_t) \equiv \pi(I_t, \eta^*_t) - v(I_t, \hat{\eta}_t, \eta^*_t, \eta^*_t) = 0. \quad (6)
\]

Then \( \eta^*_t \) is the critical type representing the PS-PBE strategies of all remaining agents.

The Proposition implies that if \( \eta^*_t \) exists and if \( \eta^*_t \in (\hat{\eta}_t, 1] \), in the PS-PBE, all agents with \( \theta \in (\hat{\eta}_t, \eta^*_t] \) will adopt and those with \( \theta \in (\eta^*_t, 1] \) will wait. Nobody adopts if \( \eta^*_t \in [\underline{\eta}, \hat{\eta}_t] \) and everybody adopts if \( \eta^*_t \in (1, \bar{\eta}] \). Next we show that for an agent of type \( \eta^*_t \) (when \( \eta^*_t \in (\hat{\eta}_t, 1] \)), his decision problem can be reduced to a two period problem.
Proposition 3 Suppose $\eta_t^*$ exists in (6) and $\eta^*_t \in [\hat{\eta}_t, 1]$. If an agent of type $\eta_t^*$ waits, 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 $\eta_t^*$ will be of the lowest type among all remaining agents. If he finds it not worthwhile to adopt 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 in period $t+1$, without anybody adopting.

The two period decision characterization of the critical type $\eta_t^*$ has some flavor of the “one-step property” of optimal decisions in Chamley and Gale (1994). They 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. On this point, our model extends Chamley and Gale (1994) to pure strategies where only one type, the critical type, is indifferent between waiting and adopting. Further, from Proposition 2, this critical type helps us pin down the equilibrium strategy of every agent.

Proposition 3 simplifies the derivation of $v(I_t, \hat{\eta}_t, \eta_t^*, \eta_t^*)$. In particular, we know that if $\eta_t^*$ exists and if $\eta_t^* > \hat{\eta}_t$,

$$v(I_t, \hat{\eta}_t, \eta_t^*, \eta_t^*) = \beta E_{t+1}[I_{t+1}, \hat{\eta}_t, \eta_t^*] \pi(I_{t+1}, \eta_t^*) 1(\pi(I_{t+1}, \eta_t^*) \geq 0),$$

where $1(\cdot)$ is an indicator function returning the value of one if its argument is true, and zero otherwise. This observation greatly simplifies our search for $\eta_t^*$.

Proposition 4 There exists a unique solution $\eta_t^* \in [\hat{\eta}_t, \tilde{\eta}]$ to (6). 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 $\eta$, with $w(I_t, \hat{\eta}_t, \eta) > 0$ and $w(I_t, \hat{\eta}_t, \tilde{\eta}) < 0$.

Note that the equilibrium strategy $\eta_t^*$ can be below the prior $\hat{\eta}_t$. Then the game ends because all remaining agent types are above $\hat{\eta}_t$. (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 in the next period is simply $\hat{\eta}_{t+1} = \eta_t^*$. 

The two period method of characterizing the PS-PBE may be useful for a range of games involving dynamic decisions with future learning. The foundation of dynamic decisions with learning is Blackwell’s theorem on comparing experiments, which is directly applicable to a single agent’s two period decision problem. Extending the theorem to more than two periods for a single decision maker has been carried out by Epstein (1980), Jones and Ostroy (1984) and Demers (1991). However, extending the theorem to multiple players in a game setting is much more complicated. Our method of identifying the critical type helps avoid the multiple period problem, and as we demonstrate below, enables us to apply some powerful results of the original Blackwell’s theorem to the adoption game.

Next we show the characteristics of the PS-PBE by showing the properties of the equilibrium critical type \( \eta_t^* \equiv \eta^*(I_t, \dot{\eta}_t) \).

**Proposition 5** \( \eta^*(I_t, \dot{\eta}_t) \) is continuous and increasing in \( I_t \) and \( \dot{\eta}_t \).

The continuity of \( \eta^*(\cdot) \) follows from the fact that \( w(I_t, \dot{\eta}_t, \eta) \) is continuous and monotone in all three arguments. Note that an infinitesimal change in \( I_t \) does not affect the “amount of information” \( I_t \) carries, which is determined by the dimension of \( I_t \), or the number of profit signals. A small change in \( I_t \) 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 \( I_t \). The information level, or the dimension of \( I_t \), can change only discretely in our model. Then we would expect discrete changes in \( \eta_t^* \) as the information level changes.

As \( I_t \) increases, the profit signals point to a higher likelihood that \( c \) is of a high value. Then adopting now becomes more attractive because it is less likely that waiting will avoid unprofitable adoption. Consequently \( \eta_t^* \) increases. Higher \( \dot{\eta}_t \) 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 \( \eta_t^* \) increases.
4 The Diffusion Paths of the New Technology

In this section we study possible diffusion paths of the new technology in the PS-PBE. Obviously, the paths are determined by the characteristics of the PS-PBE, which depends on the priors about \( \theta \) and \( e \) and other parameters of the model, and the realizations of the profit signals along an equilibrium path. For example, when \( \varepsilon \) is deterministic, informational cascade and herd behavior occur since the first adopter releases full information about \( e \). In this case, the diffusion path is an extreme version of a logistic curve: the cumulative number of adopters start at a low level in the first period, shoots up in the second period when all others who find the technology more profitable adopt, and stays constant afterwards. On the other hand, if the variance of \( \varepsilon \) is extremely high so that the profit signals provide little information about \( e \), the path is close to being flat: most agents who expect the new technology to be profitable based on the prior information will adopt in the first period. In the intermediate cases, on the paths where high values of the profit signals are realized, the diffusion occurs quickly, whereas the diffusion will be slow on paths with consistently low realizations of the profit signals.

Further, it is possible that many periods would pass before the new technology even starts to be adopted. For example, suppose \( \theta_n > \eta^*(\emptyset, \emptyset) \equiv \eta^*_1 \) for all \( n = 1, \ldots, N \). Then nobody adopts initially because every agent, based on his prior belief, expects a positive probability that there are some low cost types (those with \( \theta < \eta^*_1 \)) among the group. In the next period, agents update their beliefs, realizing that all types are above \( \eta^*_1 \): \( \hat{\eta}_2 = \eta^*_1 \). Since \( \hat{\eta}_2 > \hat{\eta}_1 \equiv \emptyset \), Proposition 5 implies that \( \eta^*_2 \) increases: now every agent is more willing to adopt given the new belief. This process of updating and increasing \( \hat{\eta}_t \) continues until period \( \tau \) when \( \theta_n < \eta^*_n \) 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 several periods and then to resume. During the stopping periods, no further information is revealed about \( e \), but information about agent types is always updated. Particularly, \( \hat{\eta}_t \) gradually “works itself up” until it surpasses the lowest type of the remaining agents.

The temporary stops highlight a key departure of our paper from Chamley and Gale (1994)
and the herd behavior literature. The fact that nobody adopts says nothing about the technology’s efficiency. Rather, temporary stops only reflect the learning process of agents about the costs or types of others. In Chamley and Gale (1994), there is really no private value, and an agent’s type (or his private signals) directly influences the technology’s efficiency, or the common value. If few people adopts, agents would judge that the technology is not worthwhile, and consequently nobody will adopt and the game ends. The adoption process can never stop temporarily. But in reality, we do observe many adoption processes that demonstrate high and low adoption rates as well as temporary stops overtime.

If information updating about the agent types is “slow” in the sense that \(\eta_t^* - \hat{\eta}_t\) is small, there may be prolonged initial delay and temporary stops. There may even be equilibrium paths lasting infinite periods, if agents have strong incentives to wait so that \(\eta_t^* - \hat{\eta}_t\) is infinitesimal. In summary, we know

**Proposition 6** *The adoption process may start late, stop temporarily, and last infinite periods.*

As long as the game does not end, the beliefs about agent types \(\hat{\eta}_t\) increases overtime. In other words, in any period before the end of the game, \(\eta_t^* > \hat{\eta}_t\) (note that \(\hat{\eta}_{t+1} = \eta_t^*\) in this case). The game ends either when all agents have adopted, or when \(\eta_t^* \leq \hat{\eta}_t\). This latter scenario is likely to arise if the signals of the new adopters in period \(t - 1\) indicate extremely low \(e\) values, i.e., if there is a “shock” of low signals. In this case, nobody adopts in this period. Further, there is no information updating about agent types because \(\hat{\eta}_{t+1} = \hat{\eta}_t > \eta_t^*\) (rather than \(\hat{\eta}_{t+1} = \eta_t^*\)). Nobody adopts not because the agents’ types are high, but because the new signals are too dismal. Thus the information set 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 if \(\eta_t^* < \hat{\eta}_t\), the game ends in period \(t\) with none of the remaining agents ever adopting. Our discussion proves the following:

**Proposition 7** *Along an equilibrium path, the adoption game ends either when all agents have adopted, or at information set \(I_t\) with \(\eta^*(I_t, \hat{\eta}_t) \leq \hat{\eta}_t\), which occurs when the new signals about the technology are sufficiently low. Before this state is reached, the prior beliefs \(\hat{\eta}_\tau\), for \(\tau < t\), increase in \(\tau\).*
We next show that a wide range of possible diffusion paths can be logistic. For simplicity, we assume that agent types and the prior are such that at least one agent adopts but not all adopt in period one. That is, $Q < \eta^*(0, \theta) < 1$, and agent types are sufficiently “dispersed” so that $\theta_n < \eta^*(0, \theta)$ for some but not all $n = 1, \ldots, N$.

Given $\{\theta_1, \ldots, \theta_N\}$, the diffusion pattern along any path is completely characterized by the evolution of $\eta^{\hat{\theta}} - \hat{\eta}_t$, which determines the number of new adopters in each period. For example, suppose the types $\{\theta_1, \ldots, \theta_N\}$ are equally spaced on $[Q, 1]$. Then the diffusion path is convex if $\eta^{\hat{\theta}} - \hat{\eta}_t$ increases overtime: the number of new adopters, which is also the increase in the cumulative number of adopters, rises as time progresses. The path is concave if $\eta^{\hat{\theta}} - \hat{\eta}_t$ decreases overtime and is logistic if it first increases and then decreases. Thus we only need to study how $\eta^{\hat{\theta}} - \hat{\eta}_t$ changes overtime to find the diffusion pattern along a particular path.

There are (uncountably) infinite number of diffusion paths, each corresponding to a particular realization of the observation errors of the adopters. Particularly, given prior $\hat{\eta}_t$, the equilibrium critical type $\eta^{\hat{\theta}}$ depends on the realized signals of the new adopters in period $t - 1$. 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. That is, there is a one-to-one correspondence between an equilibrium diffusion path and a sample selection of the observation errors of all adopters.

Now we consider a particular diffusion path (or a particular selection from the sample of observation errors). As time goes by, two changes occur to the optimal response function $\eta^*(I_t, \hat{\eta}_t)$: in terms of information about $e$, the dimension of $I_t$ becomes higher as more agents adopt (and more profit signals are observed), and in terms of beliefs about $\theta$, $\hat{\eta}_t$ increases. The two changes affect the path differently. As more signals are observed, the prior of remaining agents about $e$ becomes stronger and there will be fewer new signals as there are fewer agents left. Thus their incentive to wait for more signals decreases and they are more willing to adopt. That is, if we fix the belief $\hat{\eta}_t$, $\eta^*(I_t, \hat{\eta}) - \hat{\eta}$ increases in the dimension of $I_t$. Therefore, the fact that $\text{dim}(I_t)$ increases along the diffusion path tends to make the path convex. On the other hand, if the information set $I$ is fixed, higher $\hat{\eta}_t$ means that the remaining agents are of higher costs, and fewer of them are likely to adopt. That is, $\eta^*(I, \hat{\eta}_t) - \hat{\eta}$, which measures the length of the interval $(\hat{\eta}_t, \eta^{\hat{\theta}})$ of the new
adopter types, decreases as \( \hat{\eta} \) rises. Then as \( \hat{\eta} \) increases along the diffusion path, the number of new adopters are likely to decrease, leading to a concave diffusion path. The shape of the diffusion path depends on the relative effects of increasing \( \dim(I_t) \) and \( \hat{\eta} \), and in many cases will be logistic as the “convexity” and “concavity” effects dominate in different time periods. In what follows we formally make the above arguments.

Consider first the impacts of increasing number of signals in (or dimension of) \( I_t \), given fixed beliefs about \( \theta, \hat{\eta} \). To single out the dimension effect, we remove the influence of the magnitudes of these signals by comparing two information sets \( I^1 \) and \( I^2 \) with \( E_{eqI^1}(e) = E_{eqI^2}(e) \) and \( \dim(I^1) < \dim(I^2) \). The two sets predict the same expected payoff of adopting now, but contain different amount of information.

**Proposition 8** Consider two continuation games \((I^1, \hat{\eta})\) and \((I^2, \hat{\eta})\) with \( E_{eqI^1}(e) = E_{eqI^2}(e) \) and \( \dim(I^1) < \dim(I^2) \). Then

\[
\eta^{s1} \equiv \eta^*(I^1, \hat{\eta}) < \eta^*(I^2, \hat{\eta}) \equiv \eta^{s2},
\]

if both \( \eta^{s1} \) and \( \eta^{s2} \) are strictly above \( \hat{\eta} \).

There are two forces underlying (8). As there are fewer remaining agents under \( I^2 \), the distribution of the number of future signals shifts to the left. Further, the value of an additional signal goes down as the prior is stronger under \( I^2 \). 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.9

Proposition 8 formally establishes that if an information set contains more signals about \( e \),

---

9 Note that this result does not imply that the value of future information is concave in the number of future signals. As shown by Radner and Stiglitz (1984), the value of information may be nonconcave, particularly in the neighborhood of zero information. The value would be concave in our model if given the same prior about \( e \), the added value of two additional signals is less than double that of one additional signal. We can show this may not be true. For example, suppose the information set is such that \( \pi(I, \theta) \gg 0 \). Then one additional signal, regardless of its magnitude, may never cause \( \pi(I_1, \theta) \) to be negative. In this case, this additional unit of signal has a value of zero. But suppose two additional signals are able to overcome the prior and lead to possibilities of \( \pi(I_1, \theta) < 0 \). Then the two signals together have a positive value.
other things equal, there will be more agent types who would adopt now. This result is further strengthened if $E_{\eta^{l,1}}(e) < E_{\eta^{l,2}}(e)$ but weakened if $E_{\eta^{l,1}}(e) > E_{\eta^{l,2}}(e)$.

Next let us fix $I$, the information about $e$, and study the effects of increasing $\hat{\eta}$. We start with the baseline case where the beliefs about types are uniformly distributed: $g_0(\theta) = 1/(1-\theta)$.

**Proposition 9** Suppose the beliefs about agent types are uniformly distributed. Then for any $I \in \mathcal{I},$

$$\eta^{2*} - \eta^2 < \eta^{1*} - \eta^1,$$

if $1 \geq \hat{\eta}^2 > \hat{\eta}^1 \geq \theta$ and $\eta^{i*} > \hat{\eta}^i$, where $\eta^{i*} \equiv \eta^*(I, \hat{\eta}^i), i = 1, 2$.

The intuition in proving 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 $\eta^{2*}$ (which is higher than $\eta^{1*}$) 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 $\eta^{1*}$), which requires $\eta^{2*} - \eta^2$ to be smaller than $\eta^{1*} - \eta^1$ when the beliefs are uniformly distributed.

When the beliefs are not uniformly distributed, the result depends on the relative density of the distribution on $(\hat{\eta}^1, \eta^{1*})$ and on $(\hat{\eta}^2, \eta^{2*})$. The inequality in (9) will be strengthened if the density is higher on $(\hat{\eta}^2, \eta^{2*})$, since $\eta^{2*} - \eta^2$ has to be further smaller to reduce the number of signals in the future. The inequality will be weakened or reversed if the density is higher on $(\hat{\eta}^1, \eta^{1*})$.

Consider now a situation where $\{\theta_1, \ldots, \theta_N\}$ are equally spaced on $[\theta, 1]$. In the simple case of uniform priors about $\theta$ and constant observation errors across agents, the convexity or concavity of the diffusion path depends on the relative magnitudes of the two effects identified in Propositions 8 and 9. A logistic diffusion path is likely to result if the effects of increasing dimension of $I_t$ dominates in early periods, and those of increasing $\hat{\eta}_t$ dominates in latter periods. Showing this phenomenon involves the second order properties of $\eta^*(I, \hat{\eta})$, which in turn depends on the second order effects of information signals. As we discuss below, the second order effects of information is not clear, and we can only demonstrate some of the factors that will contribute to the logistic diffusion pattern.

We first extend Proposition 9 and argue that as $\hat{\eta}$ rises, further increases in $\hat{\eta}$ will cause even smaller increases in $\eta^*$. That is, the “degree of concavity” of the diffusion path caused by increasing $\hat{\eta}$ will be higher as $\hat{\eta}$ becomes higher.
Proposition 10 Suppose the beliefs about agent types are uniformly distributed and the information set \( I \) is given. Let \( \hat{\eta}^2 - \hat{\eta}^1 = \hat{\eta}^1 - \hat{\eta}^3 = \phi > 0 \) with \( \hat{\eta}^3 > \hat{\eta}^i \), \( \hat{\eta}^i \in [\phi, 1] \), and \( \eta^s > \hat{\eta}^i, i = 1, \ldots, 4 \). Further, let \( \eta^2 - \eta^1 = \rho_1 \) and \( \eta^s - \eta^{3s} = \rho_3 \). Then \( \rho_3 \leq \rho_1 \), or

\[
(\eta^4 - \hat{\eta}^1) - (\eta^{3s} - \hat{\eta}^3) < (\eta^2 - \hat{\eta}^2) - (\eta^1 - \hat{\eta}^1).
\] (10)

This result is due to the fact that the relative advantage of adopting now is decreasing and concave in the agent’s type, as shown in Proposition 1. That is, as \( \theta \) increases from \( \eta^1 \) to \( \eta^4 \), the advantage of adopting now decreases faster. Then to induce indifference between adopting and waiting, the required decrease in the number of signals must be higher as \( \eta^* \) increases, resulting in \( \eta^* - \hat{\eta} \) to decrease faster as \( \hat{\eta} \) increases.

We can extend Proposition 8 along a similar line if the value of additional profit signals, \( E_{I, \eta, L_1 \pi(I_1, \eta)1(\pi(I_1, \eta) \geq 0) \) in (23), is convex in the prior information measured by \( \text{dim}(I) \), where \( L \) is the number of new signals and \( I_1 \) is the information set in the next period. We know from the proof of Proposition 8 that the value of additional signals decreases in \( \text{dim}(I) \). Convexity requires that this value decreases in \( \text{dim}(I) \) at a decreasing rate. The distribution of the number of new adopters (i.e., \( L \) in equation (23)) is binomial with the parameters given by \( N - 1 - \text{dim}(I_1) \) and \( G_t(\eta) \). As \( \text{dim}(I) \) rises, the marginal change in the distribution, measured by its mean \( (N - 1 - \text{dim}(I))G_t(\eta) \) and variance \( (N - 1 - \text{dim}(I))G_t(\eta)(1 - G_t(\eta)) \), is the same regardless of the initial dimension of \( I \).

Then as the dimension of \( I \) rises, the convexity of \( E_{I, \eta, L_1 \pi(I_1, \eta)1(\pi(I_1, \eta) \geq 0) \) in \( \text{dim}(I) \) implies that further increase in the dimension leads to a smaller decrease in \( u(\cdot) \). Thus \( \eta^* \) will increase at smaller steps, and the degree of convexity of the diffusion path will decrease as more profit signals are contained in \( I_1 \). However, unlike Proposition 10, showing the convexity in \( \text{dim}(I) \) of the value of an additional signal cannot be done without further restrictions on the relevant functional forms.\(^{10}\)

Nevertheless, Propositions 8 - 10 provide arguments for the likelihood of logistic diffusion paths, in the simple case of equally spaced types, uniform priors and constant expected \( \epsilon \). Continuity of the critical type in these variables implies that a logistic path is likely in the neighborhood of these

\(^{10}\)One reason for the convexity is that the variance about \( \epsilon \) in the information set \( I \) is likely to be decreasing and convex in \( \text{dim}(I_1) \). For example, this is true if the variance is proportional to \( 1/\text{dim}(I_1) \), as in the case of \( f_0(\cdot) \) and \( \epsilon \) being normally distributed. Then the value \( E_{I, \eta, \epsilon \pi(I_1, \eta)1(\pi(I_1, \eta) \geq 0) \) is convex in \( \text{dim}(I) \), as long as it is not “too concave” in the variance about \( \epsilon \).
conditions. When these conditions change, the underlying forces discussed in these propositions still exist, pushing for a slow start and slow ending along the realized diffusion path. For example, if there is a relatively large number of “average” agent types, the logistic shape is likely to be augmented. But if there are large numbers of very high and very low types, the logistic shape will be weakened. The final diffusion pattern depends on the relative effects of the factors at work.

The pattern also depends on the specific realizations of the observation errors. There are several possibilities. (i) If consistently high signals are realized, then the path may still be logistic, but the rate of diffusion will be high. (ii) Conversely, if the adopters generate consistently low signals, the adoption rate and speed of diffusion will be low, even though the path may still be logistic. (iii) For cases where both high and low signals are generated, the path depends on how the high and low signals are realized sequentially. If early signals are consistently high and later signals are always low, the technology may quickly diffuse with sudden ending. Conversely, we may observe a slow start of the adoption process and quick diffusion afterwards. In both cases, the diffusion path may not be logistic.

5 Comparative Statics

From the last two sections, it is obvious that other things equal, if the prior $f_0(\cdot)$ about $e$ is more informative, the adoption process is likely to start earlier and the diffusion will be faster. That is, given the same $I_t$ and $\dot{\eta}$, $\hat{\eta}^*$ will be higher if $f_0(e)$ is less dispersed. 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 the prior information (Sunding and Zilberman (forthcoming)). The earlier adoption and faster diffusion of new technologies in developed countries (relative to developing countries) may be partly attributed to the more widespread availability of these services.

The diffusion will be faster if the new technology becomes more efficient or less costly. That is, $\eta^*(I, \dot{\eta})$ increases for any $I$ and $\dot{\eta}$ if the distribution of $e$, $f_0(\cdot)$, increases in the sense of first order stochastic dominance, or if $c$ becomes smaller. This conclusion is consistent with the empirical
findings that “good” technologies tend to be adopted earlier and diffused faster. This positive correlation between the extent of improvement of the new technology and the speed of diffusion is also found in Ellison and Fudenberg (1993) in a non-game setting.

In this section we investigate how the equilibrium diffusion paths respond to changes in other parameters, in particular beliefs about agent types (i.e., the variance of $\theta, \sigma^2_\theta$), learning about the technology (i.e., the variance of $\varepsilon, \sigma^2_\varepsilon$), and the number of agents in the community $N$. Within a more closely knitted community, agents may have stronger beliefs about each other, resulting in a smaller $\sigma^2_\theta$. Such a community may also have a lower $\sigma^2_\varepsilon$, but the observation error may be caused by factors exogenous to the community, such as weather variations, market shocks, or the nature of the technology itself.\footnote{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.} In our discussion, we interpret $\sigma^2_\theta$ as the strength of beliefs about types, which may be related to the degree of communication among the agents about the types, and $\sigma^2_\varepsilon$ as capturing the degree of communication about the technology.

5.1 Beliefs about Agent Types

Given an information set $I_t$ and prior information $\hat{\eta}_t$, as $\sigma^2_\theta$ increases, it is possible that an agent’s perception of the number of likely adopters in this period is either higher or lower. If the perception becomes higher, each remaining agent is more willing to wait for others to adopt and the diffusion process slows down. The diffusion becomes faster if the perception is lower. It is therefore important to first investigate how the distribution of the (perceived) number of new adopters in a period depends on $\sigma^2_\theta$.

Given an information set $I_t$ and a prior $\hat{\eta}_t$, the number of remaining agents is $N - \dim(I_t)$, and the common belief about the type of each of them is distributed according to $G_t(\cdot)$, which is $G_0(\cdot)$ conditional on $\theta > \hat{\eta}_t$ (cf. (5)). When all other remaining agents follow the strategy represented by critical type $\eta > \hat{\eta}_t$, the payoff of waiting for the critical type $\eta$, say agent $n$, is given in (19) in the Appendix, which can be written as $v(I_t, \hat{\eta}_t, \eta, \eta) = \beta E_L[I_{t+1}\hat{\eta}_t, \eta, \pi(I_t, L_t, \eta)]$, where $z(I_t, L_t, \eta) = E_{I_{t+1}|I_t, L_t, \pi(I_{t+1}, \eta)}1(\pi(I_{t+1}, \eta) \geq 0)$ is the expected payoff of waiting when the number
of new adopters (other than agent $n$) is $L_t$. Lemma 1 implies that $z(I_t, L_t, \eta)$ is monotone increasing in $L_t$.

Since the beliefs about agent types are $i.i.d.$, and each agent’s probability of adoption is believed to be $G_t(\eta)$, the number of (other) new adopters $L_t$ is believed to follow the Binomial distribution with parameters $N - 1 - \dim(I_t)$ and $G_t(\eta)$. The density function is given by

$$l(L_t) = C_{L_t}^{N - 1 - \dim(I_t)} G_t(\eta)^{L_t} (1 - G_t(\eta))^{N - 1 - \dim(I_t) - L_t}, \quad 0 \leq L_t \leq N - 1 - \dim(I_t). \quad (11)$$

Suppose the beliefs change so that $G_t(\eta)$ increases to $G'_t(\eta)$. From (11), we know $l(L_t)$ decreases for low levels of $L_t$ and increases for high levels of $L_t$, and the two density functions $l(\cdot)$ and $l'(\cdot)$ associated with $G_t(\eta)$ and $G'_t(\eta)$ cross only once. Therefore, $l'(\cdot)$ first order stochastically dominates $l(\cdot)$. Since $z(\cdot)$ increases in $L_t$, we know $E_{F^{(\cdot)}} z(I_t, L_t, \eta) > E_{F^{(\cdot)}} z(I_t, L_t, \eta)$. Therefore,

**Proposition 11** Given information set $I_t$, the expected payoff of waiting for the agent of the critical type $\eta$ is monotone increasing in the believed probability of adoption $G_t(\eta)$, where $\eta$ is the critical type followed by all other remaining agents. Consequently, the equilibrium critical type $\eta^*_t$ decreases in $G_t(\eta^*_t)$.

We do not present the formal proof since the Proposition directly follows from the above discussion. The Proposition is intuitive: since the expected payoff of waiting increases in the number of new adopters, it must be increasing in the perceived probability of adoption of each of the fixed number of remaining agents.

To identify the effects of $\sigma^2_\theta$ on the payoff of waiting $v(\cdot)$, we only need to study how an increase in $\sigma^2_\theta$ affects $G_t(\eta)$, the probability of adoption. Consider the starting period when $\eta_1 = \theta$. A mean preserving spread of $G_1(\cdot) \equiv G_0(\cdot)$ implies that $G'_1(\cdot)$ is second-order stochastically dominated by $G_1(\cdot)$. Thus $G'_1(\eta) > G_1(\eta)$ for $\eta < \bar{\theta}$ and $G'_1(\eta) < G_1(\eta)$ for $\eta > \bar{\theta}$, where $\bar{\theta}$ is the mean of $\theta \in [\underline{\theta}, 1]$. If the first period’s equilibrium critical type $\eta^*_1$ is below $\bar{\theta}$, higher uncertainty raises the perceived adoption probability $G_1(\eta^*_1)$. From Proposition 11, the expected payoff of waiting increases and $\eta^*_1$ decreases, or the adoption rate goes down. Conversely, if $\eta^*_1 > \bar{\theta}$, higher $\sigma^2_\theta$ reduces $G_1(\eta^*_1)$, leading to an increase in $\eta^*_1$ and the adoption rate. Therefore,
Proposition 12 As $\sigma^2_\theta$ increases, the initial adoption rate of the technology goes down if the first period’s equilibrium critical type $\eta_1^*$ is below the mean of $G_0(\cdot)$, $\bar{\theta}$. The initial adoption rate goes up if $\eta_1^* > \bar{\theta}$.

Again the proof follows directly from the above discussion. If $\eta_1^* < \bar{\theta}$, only agents with relatively low costs will adopt given the original uncertainty $\sigma^2_\theta$. Higher uncertainty shifts the beliefs about the distribution of agent types to the tails, leading to a higher assessment of the probabilities of low types. With higher uncertainty, each agent believes that it is more likely that other agents are of really low costs. Each expects more adopters in the first period and is more willing to wait, leading to a lower initial adoption rate.

On the other hand, if $\eta_1^* > \bar{\theta}$, only agents with really high costs will not adopt in the first period. As $\sigma^2_\theta$ increases, each agent believes higher probabilities of really high types, or a higher number of non-adopters. Thus he is less willing to wait, and the initial adoption rate increases.

In periods $t > 1$, for $\eta_t > \bar{\theta}$, we know from (5) that

$$G_t(\eta) = \frac{G_0(\eta) - G_0(\eta_t)}{1 - G_0(\eta_t)} = 1 - \frac{1 - G_0(\eta)}{1 - G_0(\eta_t)}.$$ (12)

The continuous differentiability of $g_0(\cdot)$ implies that the sign of $\partial G_t(\eta) / \partial \sigma^2_\theta$ is continuous in $\eta$ (and $\eta_t$). Thus our results in period one are likely to carry over to early periods. That is,

**Remark 1** For technologies that are adopted quickly (in particular, $\eta_1^* > \bar{\theta}$), weaker beliefs about agent types (or higher $\sigma^2_\theta$) are likely to speed up the adoption process in early periods. For technologies that are adopted more gradually, the weaker beliefs are likely to slow down the early adoption process.

If $\eta_t < \bar{\theta}$ and $\eta_t^* > \bar{\theta}$, higher $\sigma^2_\theta$ raises $G_0(\eta_t)$ and reduces $G_0(\eta_t^*)$. From (12), we know $G_t(\eta_t^*)$ decreases and $\eta_t^*$ increases. In this case, higher uncertainty about $\theta$ raises the diffusion rate. A mean-preserving spread of the distribution of beliefs implies that fewer types are expected in the middle range of the distribution. When the likely adopters in period $t$ are exactly these middle range types, each remaining agent expects fewer new adopters in period $t$ and is thus less willing to wait, increasing the diffusion rate.
If both \( \eta_t \) and \( \eta_t^* \) are higher than but close to \( \bar{\theta} \), we expect that \( \partial G_0(\eta_t^*) / \partial \sigma_0^2 < \partial G_0(\eta_t) / \partial \sigma_0^2 < 0 \): the cumulative distribution function decreases by a larger amount at the higher level of \( \theta \), \( \eta_t^* \), than at \( \eta_t \) when \( \sigma_0^2 \) rises. Combining this with the fact that \( G_0(\eta_t^*) > G_0(\eta_t) \), we can show from (12) that \( G_1(\eta_t^*) \) decreases as \( \sigma_0^2 \) increases. Thus \( \eta_t^* \) increases, so does the diffusion rate. Again, continuity of \( \partial G_1(\eta) / \partial \sigma_0^2 \) implies that

**Remark 2** For technologies that are gradually diffused, weaker beliefs about agent types (or higher \( \sigma_0^2 \)) are likely to speed up the middle part of the diffusion process.

The analysis becomes more complicated and less clear for other situations where \( \eta_t \) and \( \eta_t^* \) lie on the same side of \( \bar{\theta} \).

We should note that changes in \( \sigma_0^2 \) does not affect the adoption rate of “good technologies” in the long run. By good technologies we mean those for which all profit signals indicate high expected value of \( e \). Then whenever some agents become the lowest types (i.e., it is their turn to adopt), they always face high signals and always adopt. Higher \( \sigma_0^2 \) may slow down or speed up the adoption process, but does not change the eventual adoption rate.

However, for other technologies where both high and low signals are possible, higher \( \sigma_0^2 \) may raise or reduce the eventual adoption rate. For example, suppose with the original \( \sigma_0^2 \), the information sets are \( I_2 = \{p_1, p_2\} \) and \( I_3 = I_2 \cup \{p_3, p_1, p_5\} \): Agents 1 and 2 adopt in period one and agents 3, 4 and 5 adopt in period two. Suppose \( \sigma_0^2 \) increases to \( \sigma_0' \), and the signals are \( I_2 = \{p_1, p_2\} \) and \( I_3 = I_2 \cup \{p_3\} \). Higher \( \sigma_0^2 \) leads agents 4 and 5 to wait in period two. Then if \( p_3 \) is sufficiently low, it is possible that agents 4 and 5, after observing \( p_3 \), will never adopt, while under \( \sigma_0^2 \) they adopt before observing \( p_3 \). Further, if \( p_4 \) and \( p_5 \) are high enough, the adoption process under \( \sigma_0^2 \) will continue into later periods because the low value of \( p_3 \) is balanced by high \( p_4 \) and \( p_5 \). In this example, higher \( \sigma_0^2 \) causes premature stopping of the diffusion process and greatly reduces the long-run adoption rate.

The situation is reversed if higher \( \sigma_0^2 \) causes more agents to adopt in period two. For example, this is true if \( I_2 = \{p_1, p_2\} \) and \( I_3 = I_2 \cup \{p_3\} \) under \( \sigma_0^2 \), and \( I_2 = \{p_1, p_2\} \) and \( I_3 = I_2 \cup \{p_3, p_4, p_5\} \) under \( \sigma_0' \). In summary,
Remark 3 Changes in $\sigma^2_0$ does not affect the long-run adoption rate of extremely good technologies. However, it may raise or reduce this rate for other technologies.

5.2 Learning about the Technology

A key observation, shown in Kihlstrom (1984), is that higher $\sigma^2_0$ implies that the profit signals of the adopters carry less information about $e$ in the sense of Blackwell (1951, 1953). Then given the same prior information about $e$, these signals are less valuable and agents will have less incentive to wait for them. Therefore, in the first period, when the prior information about $e$ is fixed, $\eta^*_t$ and the adoption rate increases as $\sigma^2_0$ rises.

In later periods with $\dim(I_t) > 0$, higher $\sigma^2_0$ also reduces the prior information about $e$ contained in the set $I_t$. But when $\dim(I_t)$ is low, we expect the prior information $f_0(\cdot)$ to dominate the additional information provided by $I_t$. In this case, higher $\sigma^2_0$ may reduce the information content of additional future signals relatively more than it reduces the prior information of $f_0(\cdot)$ and $I_t$. Then $\eta^*_t$ and the adoption rate increase.\(^\text{12}\) Therefore,

Remark 4 Less communication about the technology (or higher $\sigma^2_0$) is likely to speed up the initial adoption process.

Now consider the other extreme where the information in $I_t$ dominates that in $f_0(\cdot)$, in particular where $f_0(\cdot)$ is not informative at all. We cannot obtain unambiguous results without assuming special functional forms, but it is likely that higher $\sigma^2_0$ will slow down the adoption process. To see this, suppose the functions are such that the amount of information is completely characterized by the variance of $I$, which is reduced by the number of signals in the following fashion: $\sigma^2_0 = \sigma^2_0/n$.\(^\text{13}\) Then as $\sigma^2_0$ increases, the marginal reduction in the variance from an additional signal increases. Thus future signals becomes more valuable and agents are more willing to wait. In this case, higher $\sigma^2_0$ reduces $\eta^*_t$ and the speed of diffusion.

---

\(^{12}\)This argument can also be understood from the continuity of all density and payoff functions, which implies that the response of $\eta^*_t$ to $\sigma^2_0$ is continuous. Then if $\partial \eta^*_t / \partial \sigma^2_0 > 0$ for $I_1 = \emptyset$, continuity implies that $\partial \eta^*_t / \partial \sigma^2_0$ is likely to be positive even when the dimension of $I_t$ becomes positive.

\(^{13}\)These restrictions are for illustrative purposes only. There are different measurements of the “amount of information” (for example, DeGroot (1970) discussed the entropy measure), and the variance measure is typically an over-simplification.
The analysis becomes more complicated for the intermediate case where both \( f_0(\cdot) \) and \( I_t \) contribute significantly to the (prior) information about \( e \). However, the above discussion indicates that it is likely that higher \( \sigma_\varepsilon^2 \) reduces the adoption speed at later stages of the diffusion process. That is,

**Remark 5** *Less communication about the technology is likely to slow down the diffusion process in later periods.*

Higher \( \sigma_\varepsilon^2 \) implies higher probabilities of observing extreme signals, i.e., extremely high and low \( p \) values. For a relatively good technology, more extremely high \( p \) values does not matter much since other signals are in general high (and agents will adopt when it is their turn). However, more extremely low \( p \) values can counter-balance the good signals and reduce the adoption rate. For mediocre technologies where most signals are in general low, more extremely low signals do not influence the adoption rate much. But more extremely high signals may increase the adoption rate.

**Remark 6** *Higher observation error \( \sigma_\varepsilon^2 \) may reduce the adoption rate of a relatively good technology and raise the adoption rate of a mediocre technology in the long run.*

### 5.3 Effects of Community Size \( N \)

We study the effects of higher \( N \) by considering replicas of our original game. In particular, we investigate the change in \( \eta^*(I, \hat{\eta}) \) as \( N \) doubles and the dimension of \( I \) doubles, with \( E_{\varepsilon[I]}(e) \) and \( \hat{\eta} \) unchanged. Again, in period one when \( I = \emptyset \), higher \( N \) implies that there will likely be more adopters given any \( \hat{\eta} \) (or the distribution of \( L \), the number of new adopters, shifts to the right). Thus each agent has more incentive to wait, reducing \( \eta^*_t \). The continuity argument discussed in earlier sections indicates that this may be true in more than one early periods.

**Remark 7** *Larger community size is likely to slow down the early diffusion process.*

In later periods, the effects of higher \( N \) depend on the trade-off between the larger number of existing signals and the larger number of potential future signals. In the limit, as \( N \to \infty \), information cascade occurs: since there are infinite number of adopters in period one (i.e., the
number of first period adopters in the original game is replicated infinite times), full information about $e$ is released. Then no agent will wait beyond period two to make his adoption decision. Again, the continuity argument implies that

**Remark 8** The speed of diffusion in later periods can increase as the community size gets larger.

In studying possible diffusion patterns and comparative statics, there are some cases (labeled as *Remarks*, rather than *Propositions*) where we cannot obtain definitive results. They are not due to the game itself, as the PS-PBE is completely characterized by a simple two period decision problem. These ambiguities arise from the fundamental tradeoffs in information economics, i.e., in a single decision maker’s problem. For example, how the marginal value of future signals depends on the existing and future signals drives much of the ambiguity. Our discussion in these cases helps identify the influencing factors, but the definitive results depend on the specific parameters of specific problems.\(^{14}\)

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

\(^{14}\)Due to the length of the paper, we choose not to use numerical examples to illustrate the generality of our remarks.
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 \( \theta_c \) 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 \( \gamma e \), where \( \gamma \) 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 \( \gamma \). Further, the signals of earlier adopters will have to be re-assessed, given the better information about \( \gamma \). 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 \( \gamma \). The learning process becomes nearly impossible to analyze. However, the nature of learning remains the same, and given the same beliefs about \( \gamma \) 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

It is widely believed that the experiences of early adopters of new technologies are an important information source for other people in making their adoption decisions. Communication between adopters and others thus helps promote adoption and diffusion through the information externalities or services of early adopters. However, if people are strategic, the communication opportunities can hinder diffusion when agents intentionally delay in order to learn from others. This paper studies 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 in fact slows down the early adoption process, even though it may speed up diffusion in later periods. It may raise the long-run adoption rate of a relatively good technology, but reduce that of a mediocre technology. On the other hand, stronger beliefs about each other’s types tend to speed up early adoption and slow down the diffusion later on. They do not alter the long-run adoption rate of extremely good technologies, but can either raise or reduce the rate for other kinds of technologies.

Therefore, *ex post*, after others have adopted, communication always promotes adoption. However, *ex ante*, the overall effects of communication and learning depend on the nature of the technology and the stage in the diffusion process. Promoting communication about the technology does not necessarily increase adoption and diffusion, unless it is done in such a way that nobody anticipated it.
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. 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 \((I_t, \hat{y}_t)\), the external value of a certain signal in \(I_t\), say \(p_1\), is

\[
\sum_{n \in \mathcal{N}_t} \left[ \max\{\pi(I_t, \theta_n), v(I_t, \hat{y}_t, \eta^g(I_t, \hat{y}_t), \theta_n)\} - \max\{\pi(I_t \setminus p_1, \theta_n), v(I_t \setminus p_1, \hat{y}_t, \eta^g(I_t \setminus p_1, \hat{y}_t), \theta_n)\} \right],
\]

(13)

where \(\mathcal{N}_t\) denotes the set of remaining agents. Ex ante, before \(p_1\) is realized, the external value of agent 1 adopting in period \(t - 1\) is the expectation of (13) over \(p_1\).

Another way of promoting adoption is to appropriately increase (or decrease) communication among the agents about the technology and influence beliefs about their types. For example, in the case of agricultural technologies, extension agents may organize more meetings among farmers to exchange information about a new technology in later periods (i.e., after the adoption rate has reached a certain level), but not in early periods. The opposite may be true for communication about each other’s type (or the likelihood of adoption): more intensive exchange should occur in early periods.

\hspace{1cm}{\footnotesize\(^{15}\)}An example of the traditional approach is Shampine (1998).
Proofs

Proof of Proposition 1. At each information set in the continuation games of \( I_t, I_r \in \mathcal{I}(I_t) \), with corresponding beliefs \( g_{-n}(\tau), \tau = t + 1, \ldots \), denote \( I_{t+1}^1 \) as the subset on which agent \( n \) adopts with strictly positive probability (i.e., \( s(I_{t+1}^1, \theta_n) > 0 \)), and \( I_{t+1}^0 \) as that on which agent \( n \) does not adopt (given, of course, \( s_{-n} \)). Then we know

\[
v(I_t, g_{-n}(t), s_{-n}, \theta_n) = \beta E_{I_{t+1}^1} \pi(I_{t+1}^1, \theta_n) + \beta E_{I_{t+1}^0} v(I_{t+1}^0, g_{-n}(t + 1), s_{-n}, \theta_n),
\]

where \( E_{I_{t+1}^k}, k = 0, 1 \), are the expectations over the sets \( I_{t+1}^k \) conditional on \( I_t, s_{-n}(I_t) \), and \( g_{-n}(t) \), and \( g_{-n}(t + 1) \) is updated according to (1). Functions \( \pi(\cdot) \) and \( v(\cdot) \) are the expected payoffs conditional on the specific information sets in \( I_{t+1}^1 \) or \( I_{t+1}^0 \) being realized. Note that when agent \( n \) plays mixed strategies on the set \( I_{t+1}^1 \), the expected payoffs of adopting and waiting are the same, and are represented by \( \pi(I_{t+1}^1, \theta_n) \). Similarly, \( \pi(\cdot) \) can be rewritten as

\[
\pi(I_t, \theta_n) = E_{I_{t+1}^1} \pi(I_{t+1}^1, \theta_n) + E_{I_{t+1}^0} \pi(I_{t+1}^0, \theta_n),
\]

where \( I_{t+1}^1 \) and \( I_{t+1}^0 \) are just a particular partition of \( I_t \). Then we know

\[
\pi(I_t, \theta_n) - v(I_t, g_{-n}(t), s_{-n}, \theta_n) = (1 - \beta) E_{I_{t+1}^1} \pi(I_{t+1}^1, \theta_n)
\]

\[
+ E_{I_{t+1}^0} \left[ \pi(I_{t+1}^0) - \beta v(I_{t+1}^0, g_{-n}(t + 1), s_{-n}, \theta_n) \right].
\]

We can employ similar procedures to expand the last term on the right hand side of (16). Eventually, we obtain

\[
\pi(I_t, \theta_n) - v(I_t, g_{-n}(t), s_{-n}, \theta_n) = (1 - \beta) E_{I_{t+1}^1} \pi(I_{t+1}^1, \theta_n)
\]

\[
+ E_{I_{t+1}^0} \left\{ (1 - \beta^2) E_{I_{t+2}^1} \pi(I_{t+2}^1, \theta_n) + E_{I_{t+2}^0} \left[ (1 - \beta^3) E_{I_{t+3}^1} \pi(I_{t+3}^1, \theta_n) + \cdots \right. \right.
\]

\[
\left. + E_{I_{t+1}^0} \left( (1 - \beta^{T-t}) E_{I_{t+1}^1} \pi(I_{t+1}^1, \theta_n) + E_{I_T} \pi(I_T, \theta_n) \right) \right\},
\]

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 \( \pi(I_T^0, \theta_n) \) appears in the last part of (17).

From (17), we see that the benefit of not adopting in period \( t \) arises from the potential existence of the state \( I_T^0 \). 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, \( \pi(I^0_T, \theta_n) < 0 \). Note that if \( I^0_T = \emptyset \), there is no gain of waiting and agent \( n \) simply adopts at \( I_t \).

Now we study what happens if \( \theta_n \) increases. From (1), we know that updating about other agents’ types, \( g_{-n}(\tau) \), is not affected by the change in \( \theta_n \) at any information set \( I_\tau \). If the probabilities of reaching each information set is fixed, i.e., if \( I^k_\tau \) is fixed for \( \tau = t + 1, \ldots, T \) and \( k = 0, 1 \), from the expression of \( \pi(\cdot) \) in (3), we know the expected profit \( \pi(I^k_\tau, \theta_n) \) decreases linearly. That is, the right hand side of (17) would decrease and is linear in \( \theta_n \). However, as the expected profit of adoption decreases, since the density functions of \( e \) and \( \epsilon \) are continuous and bounded away from zero, the set \( I^0_T \) becomes larger and the expected value \( \mathbb{E}_{I^0_T} \pi(I^0_T, \theta_n) \) strictly decreases for the same \( \pi(I^0_T, \theta_n) < 0 \). That is, \( \mathbb{E}_{I^0_T} \pi(I^0_T, \theta_n) \) is decreasing and concave in \( \theta_n \). Consequently, all no-adoption sets \( I^0_\tau \) become bigger. Further, this effect would again decrease the right hand side of (17). This added effect would make the right hand side concave in \( \theta_n \). That is, \( \pi(I_t, \theta_n) - v_n(I_t, g_{-n}(t), s_{-n}, \theta_n) \) decreases and is concave in \( \theta_n \).

**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 done in the Bayesian framework, we know the signals under two are more informative than under one in the sense of Blackwell (1951, 1953). The Lemma then follows from Kihlstrom (1984).

**Proof of Proposition 2.** If all other agents follow the strategy represented by \( \eta^*_t \), that is, if they adopt when their types are below \( \eta^*_t \) and wait when their types are above \( \eta^*_t \), then (6) implies that an agent of type \( \eta^*_t \) is indifferent between waiting and adopting. From Proposition 1, we know this agent will adopt if his type \( \theta \leq \eta^*_t \), and will wait if his type \( \theta > \eta^*_t \). That is, the optimal response of this agent is to adopt the same strategy represented by \( \eta^*_t \). Thus the strategy \( \eta^*_t \) represents the PS-PBE.

**Proof of Proposition 3.** Suppose the agent of type \( \eta^*_t \), say agent \( n \), waits in period \( t \). Then in
period $t+1$, the common belief about the types of all remaining agents is $\hat{\theta}_{t+1} = \hat{\eta}_t^* > \hat{\theta}_t$. That is, $\theta > \hat{\theta}_{t+1}$ for the remaining agents in period $t+1$ other than $n$.

If agent $n$ still waits in period $t+1$ along the equilibrium path, from Proposition 2, the equilibrium strategy in period $t+1$ must satisfy $\eta_{t+1}^* < \theta_n = \eta_n^*$. For all other agents, their types $\theta > \hat{\theta}_{t+1} = \eta_t^* > \eta_{t+1}^*$. That is, nobody will adopt in period $t+1$. Then the information set about $e$ is period $t+2$ is the same as in period $t+1$: $I_{t+2} = I_{t+1}$. Further, since $\eta_{t+1}^* < \hat{\theta}_{t+1}$, every agent knows that nobody should adopt, and the beliefs about agent types in period $t+2$ is also unchanged: $\hat{\theta}_{t+2} = \hat{\theta}_{t+1}$. Thus nobody adopts in period $t+2$. Continuing this argument, we know the game ends in period $t+1$ with none of the remaining agents adopting. That is, if agent $n$ does not adopt in period $t+1$, he will never adopt.

Proof of Proposition 4. We show the existence and uniqueness by demonstrating that $w(I_t, \hat{\eta}_t, \eta)$ is (i) continuous in $\eta$, (ii) decreasing in $\eta$, (iii) $w(I_t, \hat{\eta}_t, \eta) > 0$, and (iv) $w(I_t, \hat{\eta}_t, \bar{\eta}) < 0$.

Consider the decision of an agent, say $n$, of type $\eta$ when all other agents adopt the strategy represented by the critical type $\eta$. If $\eta \leq \hat{\eta}_t$, no agents adopt in this period and the game ends:

$$v(I_t, \hat{\eta}, \eta, \eta) = \beta \max\{\pi(I_t, \eta), 0\}.$$ Then

$$w(I_t, \hat{\eta}, \eta) = \pi(I_t, \eta) - v(I_t, \hat{\eta}, \eta, \eta) = \min\{(1-\beta)\pi(I_t, \eta), \pi(I_t, \eta)\}. \quad (18)$$

From (3), we know $w(I_t, \hat{\eta}_t, \eta)$ is continuous and decreasing in $\eta$ for $\eta < \hat{\eta}_t$.

If $\eta > \hat{\eta}_t$, we know the starting belief in the next period is $\hat{\theta}_{t+1} = \eta$. Let $L_t \leq N - \dim(I_t) - 1$ be the number of new adopters in period $t$ other than agent $n$. We know $L_t$ is a random variable whose distribution depends on $\eta$ 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 $\text{Prob}(\hat{\theta}_t < \theta \leq \eta) = G_0(\eta) - G_0(\hat{\eta}_t) = G_t(\eta)$ (cf. (5)). Then $L_t$ follows Binomial distribution with $G_t(\eta)$ and $N - 1 - \dim(I_t)$.

For each value of $L_t$, the starting information of the next period $I_{t+1}$ contains $I_t$ as well as the new profit signals generated by the $L_t$ new adopters. Thus, from (7), we know

$$v(I_t, \hat{\eta}_t, \eta, \eta) = \beta E_{\eta_t} E_{I_{t+1}} f_t, I_{t+1}, I_t, \pi(I_{t+1}, \eta) 1(\pi(I_{t+1}, \eta) > 0). \quad (19)$$

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

33
Similar to (15) and (16), we can decompose \( \pi(I_t, \eta) \) according to \( L_t \) and \( I_{t+1} \). Then we know

\[
w(I_t, \hat{\eta}_t, \eta) = \pi(I_t, \eta) - v(I_t, \hat{\eta}_t, \eta, \eta) = E_{L_t | I_t, \eta} E_{I_{t+1} | L_t, \eta} (1 - \beta) \pi(I_{t+1}, \eta) \mathbb{1}(\pi(I_{t+1}, \eta) \geq 0)
\]

\[
+ E_{L_t | I_t, \eta} E_{I_{t+1} | L_t, \eta} \pi(I_{t+1}, \eta) \mathbb{1}(\pi(I_{t+1}, \eta) < 0).
\]

(20)

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

\[
u(I_t, \hat{\eta}_t, \eta, \theta) = \beta E_{L_t | I_t, \eta} E_{I_{t+1} | L_t, \eta} \pi(I_{t+1}, \theta) \mathbb{1}(\pi(I_{t+1}, \theta) \geq 0).
\]

(21)

Note that \( u(\cdot) \) in (21) is not the value of waiting of type \( \theta \) when others adopt \( \eta \) (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 \( \theta \) if he decides to follow the two step decision. We use (21) only as a way of decomposing the effects of higher \( \eta \) in (20) into two parts: In one, captured by (21), the type of the critical type agent is higher. In the other, critical type strategies of other agents are higher.

Using (21), we know

\[
\pi(I_t, \theta) - u(I_t, \hat{\eta}_t, \eta, \theta) = E_{L_t | I_t, \eta} E_{I_{t+1} | L_t, \eta} (1 - \beta) \pi(I_{t+1}, \theta) \mathbb{1}(\pi(I_{t+1}, \theta) \geq 0)
\]

\[
+ E_{L_t | I_t, \eta} E_{I_{t+1} | L_t, \eta} \pi(I_{t+1}, \theta) \mathbb{1}(\pi(I_{t+1}, \theta) < 0).
\]

(22)

Then it is obvious that \( \pi(I_t, \theta) - u(I_t, \hat{\eta}_t, \eta, \theta) \) decreases in \( \theta \). Further, similar to the argument following (17) in the proof of Proposition 1, we can easily show that \( \pi(\cdot) - u(\cdot) \) is concave in \( \theta \).

Thus to show that \( w(I_t, \hat{\eta}_t, \eta) \) is decreasing in \( \eta \), we only need to show that \( u(I_t, \hat{\eta}_t, \eta, \theta) \) is increasing in \( \eta \). Lemma 1 implies that in (21), \( E_{I_{t+1} | L_t, \eta} \pi(I_{t+1}, \theta) \mathbb{1}(\pi(I_{t+1}, \theta) \geq 0) \) increases in \( L_t \).

As \( \eta \) 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 \( w(I_t, \hat{\eta}_t, \eta) > 0 \), note that \( \pi(I_t, \eta) > 0, \forall I_t \), from (3). Since \( \eta < \hat{\eta}_t \), (18) implies \( w(I_t, \hat{\eta}_t, \eta) > 0 \). Similarly, (3) implies that \( \pi(I_t, \eta) < 0, \forall I_t \). Since \( \eta > \hat{\eta}_t \), (20) implies \( w(I_t, \hat{\eta}_t, \eta) < 0 \).

**Proof of Proposition 5.** We first show that \( w(I_t, \hat{\eta}_t, \eta) \) is continuous in all three arguments.

From the proof of Proposition 4, we know \( w(\cdot) \) is continuous in \( \eta \). Assumption 2 and (2) imply
that at each $e$, the conditional density $f(e|I)$ is continuous in $I$. Thus from (3), the expected payoff $\pi(\cdot)$ is continuous in $I$. Further, from (19), we know $v(\cdot)$ is also continuous in $I_t$. Therefore, $w(\cdot)$ is continuous in $I_t$.

If $\eta^*_t \leq \hat{\eta}_t$, we know from (18) that $w(\cdot)$ is independent of, and thus continuous in $\hat{\eta}_t$. If $\eta^*_t > \hat{\eta}_t$, Assumption 2 and (19) 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 $\eta^*_t$ and the fact that $w(\cdot)$ is decreasing in $\eta$, we know that $\eta^*(\cdot)$ increases in both arguments if $w(I_t, \hat{\eta}_t, \eta)$ increases in $I_t$ and $\hat{\eta}_t$. We show the latter is true.

From (18), it is obvious that when $\eta^*_t \leq \hat{\eta}_t$, $w(\cdot)$ increases in $I_t$ and is independent of $\hat{\eta}_t$. When $\eta^*_t > \hat{\eta}_t$, $w(\cdot)$ is given by (20). As $I_t$ increases, $I_{t+1}$, which contains the signals in $I_t$, also increases. Thus $\pi(I_{t+1}, \eta)$ increases, and so does $w(\cdot)$ in (20). Then $w(\cdot)$ increases in $I_t$. $\pi(I_t, \eta)$ 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 (19), or $w(\cdot)$ increases.

So far we have shown that $\eta^*(\cdot)$ increases in its arguments. The continuity of $\eta^*(\cdot)$ then follows from the fact that $w(\cdot)$ is continuous and monotone in all three arguments.

\section*{Proof of Proposition 8.} Note that $\pi(I^i, \eta^{i_*}) > 0$, $i = 1, 2$. Otherwise, the agent of type $\eta^{i_*}$ will always want to wait, instead of being indifferent between waiting and adopting. Since $\pi(\cdot)$ is continuous in $\eta$, we know $\eta^{i_*} \in$ interior of $\{\eta \in [\hat{\eta}_t, \bar{\eta}] : \pi(I^1, \eta) = \pi(I^2, \eta) \geq 0\}$, $i = 1, 2$. We next show that when $\pi(I^1, \eta) > 0$, $w(I^1, \eta, \eta)$ is higher with $i = 2$ than with $i = 1$. The proposition then follows since $w(\cdot)$ is decreasing in $\eta$.

Similar to (19), we can write the expected payoff of delay as

$$v(I^i, \hat{\eta}_t, \eta, \eta) = \beta E_{\tilde{t}|I^i, \eta} \left[ E_{\tilde{t}^i|I^i, \tilde{t}, \eta} \pi(I^i_{\tilde{t}^i}, \eta) 1(\pi(I^i_{\tilde{t}^i}, \eta) \geq 0) \right],$$

where $I^i_{\tilde{t}^i}$ denotes possible information sets in the next period. As dim($I$) 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 $\pi(I^1, \eta) \geq 0$, the possibility of “savings” by avoiding unprofitable adoption, namely the probability of $\pi(I^2, \eta) < 0$, decreases as dim($I$) increases. That
is,
\[ \pi(I^i, \eta) - \beta E_{I^i | L} \pi(I^i_1, \eta) \mathbb{1}(\pi(I^i_1, \eta) \geq 0) \]
\[ = E_{I^i | L} \left( (1 - \beta) \pi(I^i_1, \eta) \mathbb{1}(\pi(I^i_1, \eta) \geq 0) + \pi(I^i_1, \eta) \mathbb{1}(\pi(I^i_1, \eta) < 0) \right) \]
increases as \( i \) goes from 1 to 2. Since \( \pi(I^i, \eta) \) is the same for \( i = 1 \) and \( i = 2 \), we know \( E_{I^i | L} \pi(I^i_1, \eta) \mathbb{1}(\pi(I^i_1, \eta) \geq 0) \) must decrease.

Of course, the term in the square bracket of (23) is still increasing in \( L \). Since \( N - 1 - \dim(I^2) < N - 1 - \dim(I^1) \), 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 \( \pi(I^i, \eta) \) is the same for \( i = 1 \) and \( i = 2 \), we know \( w(I^2, \tilde{\eta}, \eta) > w(I^1, \tilde{\eta}, \eta) \), or \( \eta^{2*} > \eta^{1*} \).

**Proof of Proposition 9.** Similar to (21), we use the following expression:
\[ u(I, \tilde{\eta}, \eta, \theta) = \beta E_{L | \tilde{\eta}, \eta} E_{I^1 | L} \pi(I^1, \theta) \mathbb{1}(\pi(I^1, \theta) \geq 0), \quad (24) \]
where \( I^1 \) denotes the new information sets in the next period. Suppose \( \eta^{2*} - \eta^{1*} \geq \eta^{1} - \eta^{1*} \equiv \xi \). Since the beliefs are uniformly distributed, the believed distribution of \( L \) given \( \eta^{1} \) and \( \eta^{1*} \) is the same as that given \( \tilde{\eta}^{2} \) and \( \tilde{\eta}^{2} + \xi \), and both are to the left of that given \( \eta^{2} \) and \( \eta^{2*} \). Therefore,
\[ w(I, \tilde{\eta}^{1}, \eta^{1*}) = \pi(I, \eta^{1*}) - v(I, \tilde{\eta}^{1}, \eta^{1*}, \eta^{1*}) \]
\[ = \pi(I, \eta^{1*}) - u(I, \tilde{\eta}^{2}, \eta^{2} + \xi, \eta^{1*}) \]
\[ > \pi(I, \eta^{2*}) - u(I, \tilde{\eta}^{2}, \eta^{2} + \xi, \eta^{2*}) \]
\[ > \pi(I, \eta^{2*}) - v(I, \tilde{\eta}^{2}, \eta^{2*}, \eta^{2*}) \]
\[ = w(I, \tilde{\eta}^{2}, \eta^{2*}). \]

The second equality is because the distribution of \( L \) is preserved by changing \( \tilde{\eta}^{1} \) and \( \eta^{1*} \) to \( \tilde{\eta}^{2} \) and \( \tilde{\eta}^{2} + \xi \). We change \( v(\cdot) \) to \( u(\cdot) \) to reflect the fact that \( \eta^{1*} \) is not the critical type anymore in the second line. The first inequality arises since \( \pi(I, \theta) - u(I, \tilde{\eta}, \eta, \theta) \) decreases in \( \theta \) (see (22) in the proof of Proposition 4), and \( \eta^{1*} \) is replaced by \( \eta^{2*} > \eta^{1*} \). The second inequality follows since \( \eta^{2*} > \eta^{2} + \xi \), or the distribution of \( L \) shifts to the right under \( \eta^{2*} \), which implies that \( v(\cdot) \) is higher under \( \eta^{2*} \). Note that we change \( u(\cdot) \) back to \( v(\cdot) \) since \( \eta^{2*} \) is now the critical type as \( \tilde{\eta}^{2} + \xi \) changes to \( \eta^{2*} \).
But (25) contradicts the fact that \( w(I, \eta^1, \eta^{1*}) = w(I, \eta^2, \eta^{2*}) = 0 \). Thus we must have \( \eta^{2*} - \eta^2 < \eta^{1*} - \eta^1 \).

\[ \]

**Proof of Proposition 10.** Let \( \xi_1 = \eta^{1*} - \eta^1 \) and \( \xi_3 = \eta^{3*} - \eta^3 \). Thus \( \xi_1 > \xi_3 \) from Proposition 9. Note that \( \eta^{2*} - \eta^2 = \xi_1 + \rho_1 - \phi \) and \( \eta^{4*} - \eta^4 = \xi_3 + \rho_3 - \phi \).

By definition and the fact that the beliefs about types are uniformly distributed, we know

\[
\pi(I, \eta^{1*}) - v(I, \eta^1, \eta^1 + \xi_1, \eta^{1*}) = \pi(I, \eta^{2*}) - u(I, \eta^2, \eta^2 + \xi_1, \eta^{1*})
\]

\[
= \pi(I, \eta^{2*}) - v(I, \eta^2, \eta^2 + \xi_1 + \rho_1 - \phi, \eta^{2*}).
\]

(26)

The first equality follows from the uniform distribution of the beliefs about \( \theta \), and the second is from the definition of \( \eta^{2*} \).

Suppose \( \rho_3 \geq \rho_1 \). Then the last equality in (26) implies that the following is true:

\[
\pi(I, \eta^{3*}) - u(I, \eta^3, \eta^3 + \xi_1, \eta^{3*}) > \pi(I, \eta^{1*}) - u(I, \eta^3, \eta^3 + \xi_1 + \rho_1 - \phi, \eta^{3*}).
\]

(27)

To see this, note first that changing \( \eta^2 \) to \( \eta^3 \) does not affect the equality due to the uniform distribution of the beliefs (but \( v(\cdot) \) has to be changed to \( u(\cdot) \)). Then the difference between (27) and (26) is that \( \eta^{3*} \) replaces \( \eta^{1*} \) and \( \eta^{4*} \) replaces \( \eta^{2*} \). Even if \( \rho_3 = \rho_1 \), the concavity of \( \pi(I, \theta) - u(I, \hat{\eta}, \eta, \theta) \) in \( \theta \) ((22) in Proposition 4) implies (27): if \( \eta \) increases by the same amount \( \rho_3 = \rho_1 \), the decrease in \( \pi(\cdot) - u(\cdot) \) should be higher when \( \eta \) starts at \( \eta^{3*} \) than at a lower value of \( \eta^{1*} \). The inequality is further strengthened if \( \rho_3 > \rho_1 \) since \( \eta^{4*} \) would then be even higher.

Let \( \xi_1 \) decreases to \( \xi_3 \) in (27). Since \( \eta^{4*} > \eta^{3*} \), we know that in (24), \( E_{I\mid \eta, \eta} \pi(I^1, \eta^{4*})1(\pi(I^1, \eta^{4*}) \geq 0) < E_{I\mid \eta, \eta} \pi(I^1, \eta^{3*})1(\pi(I^1, \eta^{3*}) \geq 0) \). Then as \( \xi_1 \) decreases, \( u(\cdot) \) will decrease more under \( \eta^{3*} \) than under \( \eta^{1*} \). That is, the left hand side of (27) will increase more than the right hand side:

\[
\pi(I, \eta^{3*}) - u(I, \eta^3, \eta^3 + \xi_3, \eta^{3*}) > \pi(I, \eta^{4*}) - u(I, \eta^3, \eta^3 + \xi_3 + \rho_1 - \phi, \eta^{4*}).
\]

(28)

However, since \( \rho_3 \geq \rho_1 \) and \( u(I, \hat{\eta}, \eta, \theta) \) increases in \( \eta \), we know that (28) implies

\[
0 = \pi(I, \eta^{3*}) - v(I, \eta^3, \eta^3 + \xi_3, \eta^{3*}) > \pi(I, \eta^{4*}) - u(I, \eta^3, \eta^3 + \xi_3 + \rho_3 - \phi, \eta^{4*})
\]

\[
\equiv \pi(I, \eta^{4*}) - v(I, \eta^3, \eta^3, \eta^{4*}) = 0,
\]

which is a contradiction. Thus we must have \( \rho_3 < \rho_1 \).
References


Shampine, Allan, “Compensating for Information Externalities in Technology Diffusion Models,”

Sunding, David and David Zilberman, “The Agricultural Innovation Process: Research and
Technology Adoption in a Changing Agricultural Sector,” in Richard Just and Gordon Rausser,
eds., Handbook of Agricultural Economics, forthcoming.

Vettas, Nikolaos, “Demand and Supply in New Markets: Diffusion with Bilateral Learning,”

Zhang, Jianbo, “Strategic Delay and the Onset of Investment Cascades,” RAND Journal of