Technological Diffusion with Social Learning

Sandeep Kapur

The Journal of Industrial Economics, Vol. 43, No. 2 (Jun., 1995), 173-195.

Stable URL:
http://links.jstor.org/sici?sici=0022-1821%28199506%2943%3A2%3C173%3ATDWSL%3E2.0.CO%3B2-J

The Journal of Industrial Economics is currently published by Blackwell Publishing.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/black.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Thu Jun 30 14:37:49 2005

THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821 $2.00
Volume XLIII June 1995 No. 2

TECHNOLOGICAL DIFFUSION WITH SOCIAL LEARNING*

SANDEEP KAPUR

This paper attributes the slow diffusion of innovations to an informational externality in the adoption process. The profitability of new technologies is uncertain but firms can learn progressively through observing the adoption experience of others. Given this prospect of social learning, every firm would prefer that other firms adopt before it does. In the absence of explicit coordination, the firms could end up in a sequence of waiting contests. This results in staggered adoptions even when all firms are ex-ante identical. The pace of diffusion is determined endogenously, and shown to depend on the characteristics of the innovation.

I. INTRODUCTION

EMPIRICAL studies of technological diffusion suggest that the spread of new technologies is usually a gradual process, with considerable lags between successive adoptions of an innovation by various users. This paper attempts to relate the lags in the sequence of adoptions to an informational externality in the adoption process. When an innovation arrives each firm is uncertain, in a probabilistic sense, of its profitability but can progressively learn more about it through observing the adoption experience of other firms. Given this prospect of 'social learning', every firm prefers that other firms adopt before it does, because this enables a better-informed adoption decision. In the absence of any explicit coordination, the firms could end up in a sequence of waiting contests. In each contest the set of firms that have not adopted previously must choose adoption times and, in equilibrium, each firm could end up randomising over adoption times. The following pattern emerges. After one or more firms adopt, the remaining firms use the information gained through the adoption to revise their beliefs about the technology. They then engage in another waiting contest to see which of their member should adopt next, which generates some more information, and so on. The analysis of the sequence of waiting contests allows us to relate the pace of diffusion to some fundamental characteristics of the innovation, and to make some conjectures about the likely pattern of diffusion. We show that the pace of adoptions tends to rise over time if experience with the technology reinforces confidence in its profitability, while a run of adverse experience may arrest the process of diffusion altogether.

Uncertainty about the profitability of the innovation is here confined to

* I am grateful to Ashish Arora, Jayasri Dutta, Frank Hahn and Marco Mariotti for comments and to an anonymous referee for pointing out an error in an earlier version.

© Basil Blackwell Ltd. 1995, 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.

the idea that the cost of adopting the innovation (or more generally, the cost of switching from an existing technology to the new one) is uncertain. Indeed, new process technologies typically require considerable firm-specific debugging of their routines before they are truly operational and, ex-ante, the total cost of debugging (including the production loss due to down-time) is often uncertain. Every time some firm adopts the innovation, other firms can improve their estimates of the true cost of adoption. Clearly, confining the uncertainty to adoption costs alone is restrictive but makes our analysis tractable. First, it allows us to posit that the post-adoption information becomes available soon after an adoption, which would not be the case if the uncertainty related to, say, the long term profitability of the technology. Second, it ensures that new information arrives only when an adoption occurs, so that the waiting contests between successive adoptions can be modelled as *stationary* games. This would not be appropriate if, say, recurrent information from firms using the new technology continued to alter the potential adopters' beliefs, even in the absence of any further adoptions.

In our model, each waiting contest involves an adoption timing game among the potential adopters left at that stage; the contest ends as soon as at least one of them adopts. A priori, simultaneous adoptions are not ruled out. The firm(s) that adopt the innovation provide a noisy signal about the adoption cost to the remaining firms and, apart from this, they have no bearing on the future course of adoptions. The remaining firms use this information to revise their beliefs in a Bayesian manner and then engage in the next waiting contest; that ends with the next adoption(s), and so on. Adoption is assumed to be irreversible. The sequence of waiting contests continues until either all firms have adopted, or the process attains a state where no firm ever adopts, arresting the process of diffusion irreversibly. To restrict the complexity of the underlying adoption timing game, we consider the symmetric Markov Perfect Equilibria of the sequence of waiting contests. Under fairly mild restrictions, the equilibrium Markov strategies involve randomisation over adoption times. To understand this, note that the benefit from delaying adoption lies in the possibility of learning from the experience of other firms, provided they adopt first. However, because future profits are discounted, delaying adoption in the hope that some other firm will adopt first is costly. At the mixed strategy equilibrium, the benefits and costs of delaying are equal so that each firm is indifferent between various adoption times, and therefore prepared to randomise over them. We use a discrete time formulation for the waiting contests; the limiting case where the decision interval is arbitrarily small enables a reasonably simple formulation for the equilibrium randomisation. This allows us to compute the expected duration of a typical waiting contest, and thereby comment on the pace of diffusion. We also outline an argument that suggests that the coordination problem central to this analysis may lead to excessive delay in the diffusion of new technologies.

© Basil Blackwell Ltd. 1995.

Our analysis allows us to make some useful conjectures about the likely shape of diffusion curves. It is a well-noted empirical regularity in the literature on technological diffusion that adoptions of an innovation tends to be staggered, with some firms adopting it well before others. (Indeed, the time-profile of diffusion curves is often S-shaped, as demonstrated by Griliches [1957] for the spread of hybrid corn, and by Mansfield [1968] for a range of industries.) The staggered pattern of diffusion has usually been explained by appealing to the intrinsic heterogeneity between firms, as in Davies [1979]. This explanation relies on the fact that the gain to adopting a new technology varies with a firm's attributes such as size, vintage of existing fixed capital, past experience with related technologies, extent of diversification, etc. The precise configuration of these attributes determines, for each firm, a reservation price for the acquisition of the technology. Quite possibly the distribution of reservation prices might be such that not all firms find it advantageous to adopt a technology as soon as it arrives. Diffusion results from changes in the underlying characteristics of the new technology which alters the reservation prices over time. For instance, where the new technology is indivisible, initially it may be confined to only large firms whose scale of operation justifies its adoption. Subsequently, progressive reduction in acquisition costs may make the technology profitable for the smaller firms as well. If so, the shape of the diffusion path reflects the distribution, in the population of potential adopters, of the attributes that affect the gain from adoption: an S-shaped diffusion path might emerge from, say, a normal distribution of characteristics, since the cumulative normal is S-shaped.[1] While explanations of diffusion based on such heterogeneity of users are very convincing, and perhaps true for most technologies, we demonstrate the possibility of staggered adoptions even when all users are ex-ante *identical*. Our model is not unique in this respect. Other game-theoretic models have demonstrated that identical firms could end up adopting at different times - see Dasgupta [1986], and also the survey by Reinganum [1989] - but these analyses do not say much about the size of the time lags between successive adoptions. Our approach allows us to relate the lags to perceptions about the technology, and hence comment on how changing perceptions affect the pace of diffusion. This enables us to construct crude diffusion curves. We find that for good technologies (good in a sense made precise below), the early pattern of diffusion suggested by our model is not unlike the initial, rising part of an S-shaped curve.

[1] The S-shaped pattern was also explained through 'epidemic-type' models which argued that all firms do not adopt an innovation immediately because they are not all aware of its existence. If awareness spreads contagiously, through random matching between informed and uninformed agents, a logistic time path for adoptions would emerge. However, this is not very convincing for industries in which advertising and other institutional arrangements provide prompt information about the availability of the technology.

© Basil Blackwell Ltd. 1995.

The idea that firms might delay the adoption of a technology to gather more information about it is not entirely novel - it can be found in Rosenberg [1976], Balcer and Lippman [1984], Bhattacharya, Chatterjee and Samuelson [1986] and Jensen [1982], among others. However, these models tend to be decision-theoretic and ignore strategic issues. For instance, Balcer and Lippman [1984] consider a situation in which the technology improves exogenously over time, and a firm must choose between adopting the current best-practice technology and waiting for an improved version to arrive: the adoption timing problem is one of inter-temporal optimisation rather than one of strategic choice. In Jensen's model each firm chooses its adoption time optimally in relation to an exogenously-given profile of expected information flows, and diffusion is explained in terms of industry-wide differences in prior beliefs about the innovation; this replaces heterogeneity in firm characteristics with heterogeneity in beliefs as the principal explanation for different adoption times. Our model abstracts from heterogeneity completely, by imposing symmetry in beliefs as well as in all firm characteristics. Further, in our model, the arrival of information is endogenous to the process of diffusion and is a crucial aspect of the strategic interaction between firms.

While considering the positive informational externality in the adoption process, our model ignores other externalities, most notably those that arise through product-market interactions, that make the adoption decision a matter of strategic choice. (See Fudenberg and Tirole [1985] and Reinganum [1989] for models along these lines). Typically these other externalities imply an early-mover advantage in the adoption timing game while, in contrast, in our model the advantage lies with the late-movers in the adoption sequence. In terms of the categories proposed by Dasgupta [1988], ours is a waiting game as opposed to a race. There are previous analyses in which the strategic interaction has the characteristics of a waiting game. Reinganum [1985] models a late-mover advantage in the research and development process as a waiting game, and even more closely to our approach, Mariotti [1992] considers a waiting game based on an informational externality. However, both these papers consider only a *single* waiting contest rather than a *sequence of* waiting contests as in our model. With just a single waiting contest, once the first move is made by one player *all* others follow immediately. Hence these models do explain the initial delay in the adoption of a good technology but do not really explain, in a multi-firm context, why adoptions are staggered. More generally, the influence of learning on diffusion of technologies has also been considered by Besley and Case [1993] and by Ellison and Fundenberg [1993] in slightly different contexts; in fact, the term *social learning* has been borrowed from the latter. Lastly, the idea of endogenous revelation of information resulting in a late-mover advantage has been explored in contexts other than technology choice; Chamley and Gale [1994] model investment delay as an N-player game with an informational externality, but their concerns and model-structure are different from ours.

© Basil Blackwell Ltd. 1995.

The scheme of the paper is as follows. Section II sets up the formal model and describes the equilibrium in the underlying game. Section III explores the implications of this analysis for the diffusion path. Section IV outlines some extensions and concludes.

II. A MODEL OF DIFFUSION

A new, improved technology arrives in an industry with N identical firms. For each firm, the private incremental gain from switching to the new technology is assumed to be independent of other firms' technology choices; the present discounted value of the stream of future gains is known to be $R > 0$. Adoption entails sunk-cost c, uniform across firms but, ex-ante, its true value is unknown: for simplicity we assume that it could take one of two possible values, $\underline{c}$ or $\bar{c}$. Both R and c are time-invariant and, therefore, so is the net gain from adoption, $\theta = R - c$. We write $\theta(\bar{c}) = \theta_l$ and $\theta(\underline{c}) = \theta_h$, and it is assumed that $\theta_l < 0 < \theta_h$. In other words, adoption is profitable if and only if the true cost of adoption turns out to be $\underline{c}$. If so, we say that the innovation proves to be 'good'.

When the innovation arrives, the firms have exogenously-given beliefs about the true value of c. All firms are assumed to have the same prior beliefs. Subsequently, as some firms adopt the innovation, the beliefs of the others who remain to adopt (i.e., the potential adopters) are progressively modified in the light of the adopters' experience. Formally speaking, whenever some firm(s) adopt the innovation, the rest observe, possibly with some lag, a signal that is imperfectly correlated with the true cost of adoption. At any stage all potential adopters receive identical signals, and have identical revision mechanisms. This implies that all potential adopters have symmetric beliefs at any time. Learning through other firms' adoptions, or what we term as social learning, is the only channel for new information about the technology in this model. And since firms earn only through observing others' adoptions, learning is endogenous to the process of diffusion.

We measure time from the date that the innovation arrives. At any time $t \geqslant 0$, let p_t be the current probabilistic belief, common for all potential adopters, that the innovation is good, where $0 < p_t < 1$. Given the structure of our learning mechanism, this belief changes only in response to the signals that result from others' adoptions and is otherwise time-invariant. This permits the following construction. A *stage in the diffusion* of the innovation (henceforth, a *stage*) refers to the time-interval over which p_t remains unchanged. When beliefs change in response to a new signal, the diffusion process is said to move from one stage to the next one. The duration of the stages is determined endogenously in this model, a feature that allows us to study the speed of diffusion. We index the stages as follows. Let y_t be the cumulative number of the firms that have adopted the technology before time t, so that $n_t = N - y_t$ is the residual number of potential adopters at

© Basil Blackwell Ltd. 1995.

that time. A stage, and the associated value of p, are indexed by n, the number of potential adopters remaining at the start of that stage. That is, the n-th stage is described as (n, p_n).

Let the n-th stage begin at time t_n and suppose that at some time $\tau_n \geqslant t_n$, exactly k firms adopt simultaneously, where $0 < k < n$. The other $n - k$ firms observe the acts of adoption immediately, and a little later, at time $\tau_n + \delta$, they receive a signal. Here, δ is to be understood as an information lag. Let S_{nk} be the finite set of stage-n signals conditional on k firms adopting simultaneously, with typical element s. The firms' revision mechanism is Bayesian: given the current belief p_n at stage n, the updated belief on receiving a signal s is

$$(1) \qquad \tilde{p}(s, p_n) = \frac{\sigma_h(s)p_n}{\sigma_h(s)p_n + \sigma_l(s)(1 - p_n)}$$

where $\sigma_h(s)$ is the likelihood of receiving signal s conditional on $\theta = \theta_h$, and likewise $\sigma_l(s)$ for $\theta = \theta_l$. Together the set S_{nk} and the associated σ_i's define an information structure in the sense of Blackwell [1951]. [Abusing the notation slightly, we used the symbol S to represent the information structure and sometimes just the associated set of signals; the context makes clear which one we refer to.] In order to impose some consistency in the revision mechanism across the stages, we posit that $S_{nk} = S_k$ for all n. This implies that the set of potential signals and the conditional probabilities σ_i's do not vary with the stage index n, but we do allow them to depend on the number of adoptions k. The latter feature is important: if a large number of firms adopt simultaneously, we expect the resulting signal to be relatively more informative. We will return to this issue later but for the moment it completes the description of the learning mechanism.

The possibility of learning creates an informational externality. Firms that adopt at a later stage can condition their adoption decision on the information gained from the previous adoptions and, for that reason, late-adopters can expect to make better-informed decisions. Given this externality, it may not be optimal for a firm to adopt the innovation immediately even if the firm is risk-neutral and the expected profitability is positive. Rather, *each* firm would prefer that the other firms adopt first and hope to follow in their slip-stream: we have an instance of what Dasgupta [1988] calls a waiting game. We model this situation as a special kind of waiting game, namely the n-player War of Attrition. This is a generalisation of the two-player game introduced by Maynard Smith [1974], and also discussed by Kapur [1994].

The game is as follows. Starting at $t = 0$, each firm must choose some adoption time $t \in \{0, \Delta, 2\Delta, \ldots\}$. A decision never to adopt amounts to choosing an infinitely large adoption time. Notice that we model the decision process as being discrete, with Δ as the time interval between adjacent decision nodes. We refer to Δ as the reaction lag: it also measures the lag between the arrival

© Basil Blackwell Ltd. 1995.

of a signal and the earliest informed response to it. The limiting case as Δ is infinitesimally small will be of particular interest for our results. At any time the action set for a firm is $A = \{\text{wait, adopt}\}$ if the firm has not adopted before then, and otherwise is the trivial action 'do nothing'. The observed history at any time t comprises the actions taken by all firms until $t - \Delta$, and any signals received until then. For the sake of tractability, we confine our attention to Markov Perfect Equilibria (MPE) of this game. The details of this solution concept may be found in Fudenberg and Tirole [1991], but a brief description is helpful at this juncture. In some games the previous history might influence current and future play only through its effect on some 'state' variables. The state is a summary of the payoff-relevant history, where what is and what is not relevant depends, of course, on the chosen specification of the game. For instance, in the learning mechanism used here, the beliefs at any stage incorporate the information from all previous signals, so that each firm might condition its actions on the current state of beliefs rather than on the entire history (i.e., evolution) of beliefs up to that time. If so, we can define strategies as maps from the set of states to the set of actions, and such strategies are known as Markov strategies; Markov Perfect Equilibria are defined on the resultant strategy space.

In this model, the payoff-relevant aspects of the history at t are (i) the number of firms y_t that have adopted before time t, since this determines the residual number n that are still active in the game; (ii) the current value of p at any time, which incorporates prior beliefs and all the information provided by adoptions that occurred until $t - \delta$; (iii) the pattern of adoptions in the interval $(t - \delta, t)$ as this determines whether or not any signals are anticipated (with probability one) in the future interval $(t, t + \delta)$; and lastly, (iv) the elapsed time t; firms discount future returns, so that any given level of profits is more valuable at an earlier date. These four aspects together define the state at any time. Strategies for each firm can be defined as mappings from the set of all possible states to the space of probability distributions over the set of feasible actions.

It is analytically convenient, however, to decompose this game into a sequence of *stage-games*. Within each stage-game, say the n-th one, the beliefs p_n are fixed, so that the other aspects provide a complete description of the relevant history in that stage-game. Loosely speaking, the stage-games correspond to subgames in the original game, and if we can specify Nash equilibrium strategies for each stage-game, these, taken over all possible stage-games, would constitute a MPE of the original game. The relevant history within the stage-game is given by n, t and the pattern of adoptions in the interval $(t - \delta, t)$. To formalise the latter, let m be the smallest integer such that $m\Delta \geqslant \delta$. Loosely speaking m denotes the number of decision nodes that a firm would have to wait after some firm adopts to receive the signal. The pattern of adoptions in the interval $[t - m\Delta, t - \Delta]$ can be expressed as a $1 \times m$ vector

© Basil Blackwell Ltd. 1995.

$$v_t = (v^{t-m\Delta}, \ldots, v^{t-2\Delta}, v^{t-\Delta}), \qquad \text{for all } t \geqslant t_n,$$

where, $v^{t-j\Delta}$ is the number of adoptions that occurred at $t - j\Delta$. The vector v_t indicates the number of adoptions that occurred at each decision-node in the interval from $t - \delta - \Delta$ to $t - \Delta$, and we refer to it as the 'recent adoption history at t'. Within the stage-game, behaviour strategies are defined as follows. The n-th stage-game begins at t_n, and each of the n firms in this stage-game must choose an adoption time $t \in \{t_n, t_n + \Delta, t_n + 2\Delta, \ldots\}$. For firms that have adopted before time t, the only feasible action is 'do nothing'. For any firm i that has not adopted before time t, let $\beta_i(t, v_t, n; p_n)$ denote the probability of adoption (in period t) if the state is given by (t, v_t, n). The stage-contingent mapping $\beta_i(\cdot\,; p_n)$ from the set of states to the set of probability distributions over {adopt, wait} represents a behaviour strategy for firm i in this stage-game.

The firms maximise expected profits. Let $\pi(p) = p\theta_h + (1 - p)\theta_l$ denote the expected value of the net gain from adoption, given beliefs p. Now consider the expected profit if the firm can condition the adoption decision on some additional information. Suppose a firm starts with a prior p, and adopts if and only if the signal $s \in S_k$ reveals the expected profits to be non-negative. We define the information-augmented expected gain from adoption as

$$\bar{\pi}(p|S_k) = \sum_{s \in S_k} \Pr(s) \operatorname{Max}[\pi(\tilde{p}(s, p)), 0]. \tag{2}$$

Here $\Pr(s) = \sigma_h(s)p + \sigma_l(s)(1 - p)$ is the ex-ante probability, given p, of receiving the signal s. Recall that $\tilde{p}$ is the updated belief if signal s is received. In general, conditioning the decision on additional information cannot make the expected outcome any worse, so we have $\bar{\pi}(p|S_k) \geqslant \pi(p)$. Note that S_k refers to the information structure associated with k simultaneous adoptions. As stated earlier, it is entirely reasonable to posit that the informativeness of signals is increasing in k. More formally, we expect the information structure S_{k+1} to be at least as fine as S_k. It follows, from Blackwell's [1951] celebrated result, that we must have $\bar{\pi}(p|S_{k+1}) \geqslant \bar{\pi}(p|S_k)$. In words, we do not expect to be worse off with more information. To make the notation less cumbersome, we write $\pi(p_n)$ as π_n, and $\bar{\pi}(p_n|S_k)$ as $\bar{\pi}_{|n,k}$.

Note that the ordering, as above, of expected payoffs according to the availability of information is weak. For the stage-game to be a War of Attrition, it must be the case that it is *strictly* better to follow than lead in the adoption sequence. The ordering of expected payoffs with no information and with some information would be strict, that is we would have $\bar{\pi}_{|n,k} > \pi_n$, if the anticipated information S_k is of consequence to the adoption decision. Information is said to be of consequence if it affects the decisions in a non-trivial manner. This requires that there be some signal in S_k, with a non-trivial probability of arrival, which would induce the firm to behave differently than if no information was available. We assume that the

© Basil Blackwell Ltd. 1995.

anticipated information is significant in this sense. However, we need to strengthen this even further. Given that the earliest that a firm can take an informed decision entails some lag (it takes δ after any adoption for the signal to arrive, and another Δ before the firm can react to it) we must assume that these lags are small enough not to overwhelm the advantage of an informed decision. If $\rho > 0$ is the rate of interest, we restrict the validity of this model to situations where

Condition 1: $e^{-\rho(\delta+\Delta)}\bar{\pi}_{|n,k} > \pi_n$ for all $0 < k < n$.

This condition stipulates that the anticipated information is of sufficient consequence, and that the cost of waiting is sufficiently small, so that once an adoption occurs, it always makes sense to wait for the signal that would result. For given ρ, S_k and p_n, this condition imposes an upper bound on the magnitude of the lag $\delta + \Delta$. Alternatively, it can be shown that for given δ, Δ, ρ, and S_k, it imposes an upper bound on the values of p_n. On the whole, the condition is more likely to be valid in the early stages of the diffusion of technologies where, typically, the initial scepticism about the innovation is combined with the possibilities of substantial learning.

Let $G_\Delta(n, p_n)$ denote the n-th stage-game, parametrised by the size of the decision interval Δ, where $n > 1$. We seek a characterisation for the symmetric Nash equilibria of this game in behaviourally-mixed strategies.[2] Let $(\beta_\Delta^*, \beta_\Delta^*, \ldots, \beta_\Delta^*)$ be the symmetric Nash equilibrium, if it exists, for $G_\Delta(n, p_n)$, where the stage-contingent mapping $\beta_\Delta^* = \beta_\Delta^*(t, v_t, n; p_n)$ specifies the equilibrium adoption probability for a typical firm, given the state (t, v_t, n). Let α denote the following limit, whenever this limit exists:

$$\alpha(t, v_t, n; p_n) = \lim_{\Delta \to 0} \frac{\beta_\Delta^*(t, v_t, n; p_n)}{\Delta}.$$

We argue below that α could be viewed as the instantaneous adoption intensity. However, we first show

Proposition 1. Given Condition 1, as $\Delta \to 0$, the symmetric Nash equilibrium of $G_\Delta(n, p_n)$ is given as $[\alpha(n, p_n), \alpha(n, p_n), \ldots, \alpha(n, p_n)]$, such that

$$\alpha = \begin{cases} \hat{\alpha}(n, p_n) & \text{if } \pi_n \geqslant 0 \text{ and } v_t = 0; \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{where } \hat{\alpha}(n, p_n) \equiv \frac{1}{(n-1)} \frac{\rho \pi_n}{e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n}.$$

[2] There exists an asymmetric equilibrium in which one of the n firms chooses to adopt in this state and the rest of the firms choose to not adopt. This case is degenerate in the sense that in equilibrium exactly one firm adopts immediately, and is not very interesting for the analysis of diffusion. Also, given the intrinsic symmetry of the problem, they are not necessarily more convincing.

© Basil Blackwell Ltd. 1995.

Formal proofs for all propositions are in the Appendix, but some intuition for this one can be furnished here. The proposition claims that the equilibrium strategy for each firm entails either a zero adoption intensity or a stationary adoption intensity $\hat{\alpha}$, the choice between these two depending on the expected profitability (given current beliefs) and on the recent adoption history v_t. To see why it must be so, first consider the two sub-cases that lead to the adoption intensity being zero in equilibrium. Suppose that p_n is so low that π_n is negative: it is clearly sub-optimal for any firm to adopt, since adoption leads to a lower return than does waiting, which guarantees at least a nil return. Hence, each firm chooses $\beta_t^* = 0$ for all (t, v_t, n) whenever π_n is negative. And since this choice is independent of Δ, in the limit too, we must have $\alpha = 0$. No firm ever adopts and, intuitively speaking, pessimism about the innovation halts the process of diffusion altogether. Next, suppose $\pi_n \geqslant 0$ but $v_t \neq 0$. The former implies that adoption is profitable, but the latter indicates that at least one firm adopted the innovation recently. A signal is then surely expected (i.e., with probability one) in the near future, allowing the firm to make an informed decision by $t + \delta$. Given Condition 1, it is optimal for the firm to wait for that signal. Once again, $\alpha = 0$ though in this case choice reflects, not pessimism about the technology, but the desire to incorporate more information in the adoption decision. Since this holds for every potential adopter, the aggregate consequence in this case is that each adoption is followed by an 'information gathering phase' of duration δ.

Now consider the interesting case where $\pi_n \geqslant 0$ (so that the innovation is profitable) and $v_t = 0$ (no information is expected before $t + \delta$). Note that as long as no firm adopts, the process remains in the state $v_t = 0$ so that the game is stationary—if t is reached, the game looks the same as it did at $t - \Delta$. The symmetric mixed-strategy equilibrium then involves stationary strategies, and each firm chooses a constant adoption probability in each decision period. The appendix derives the equilibrium condition for the constant adoption probability, and then evaluates its limit as the decision interval $\Delta \to 0$ to obtain the equilibrium adoption intensity $\hat{\alpha}$. Condition 1 guarantees that $\hat{\alpha}$ is well defined and, indeed, positive for any $\pi_n > 0$.

For any firm, the choice of a positive (and finite) adoption intensity in the stage-game amounts, in effect, to a randomisation over the adoption decision. To see this, we first define $\alpha_\Delta = \beta_\Delta^*/\Delta$, so that $\alpha = \mathrm{Lim}_{\Delta \to 0}\, \alpha_\Delta$. Since the time interval between adjacent decision nodes is Δ, there would be $\kappa = t/\Delta$ decision nodes in any given interval 0 to t. If the probability of adoption at any decision node is β_Δ^*, the probability of not adopting at that decision node is $(1 - \beta_\Delta^*)$, and given the stationarity of the stage-game, the probability of not adopting by t equals

$$(1 - \beta_\Delta^*)^\kappa = [1 - \Delta\alpha_\Delta]^{t/\Delta}$$
$$\approx \exp(-\alpha t) \qquad \text{for small } \Delta.$$

© Basil Blackwell Ltd. 1995.

Therefore, in the limiting case described here, the cumulative probability of adoption by t equals $1 - e^{-\alpha t}$, which is the cumulative of the (negative) exponential distribution. In other words, in equilibrium each firm chooses an exponential distribution over adoption times. For firms to be prepared to randomise over adoption times, it must be the case that they are indifferent between adoption and waiting; those who wait can hope to make a more informed decision, but the gain from this is wiped out by the delay that waiting for information entails. In a sense, firms vie with each other to be followers rather than leaders in the adoption sequence, so it is not altogether misleading to refer to the choice of a positive adoption intensity $\hat{\alpha}$ as defining the 'strategic phase' with the stage-game.

Notice that $\hat{\alpha}$ depends on the characteristics of the innovation in a plausible manner. The term $\rho\pi_n$ in the expression for $\hat{\alpha}$ represents the opportunity cost (measured in flow terms) of delaying adoption: the higher is this cost, the greater must be the eagerness to adopt or, in terms of this model, the higher is the adoption intensity. The expression $[e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n]$ in the denominator represents the expected incremental gain from waiting for the next signal. The higher is this gain, or equivalently, the greater is the significance of the anticipated information, the lower is the equilibrium adoption intensity.[3] In fact, since both the numerator and the denominator vary with p_n, we can relate $\hat{\alpha}$ and p_n directly. Assume that Condition 1 continues to hold so that $\hat{\alpha}$ is non-negative. Then

Proposition 2. $\hat{\alpha}(n, p_n)$ is increasing in p_n.

The intuition for this proposition is easy to grasp. The greater is the value of p_n, the greater is the confidence in the new technology. This implies that the expected gain from immediate adoption is high, and furthermore given the Bayesian revision mechanism, it suggests that the potential gain from waiting for further information is low: both these factors translate into a higher adoption intensity.

III. THE AGGREGATE ADOPTION INTENSITY AND EXPECTED DIFFUSION CURVES

For the rest of the paper we continue to assume that Condition 1 is valid, and confine our attention to the limiting case described in Proposition 1. For that case, the form of the solution suggests that adoptions tend to be

[3] However, one must be cautious in pushing this argument too far. Strictly speaking, in the mixed-strategy equilibrium, the adoption intensity of any firm depends on the profitability of the rivals' adoption. In the case where all firms are identical, this creates no complications. In the asymmetric case, it leads to the counter-intuitive result that the firm that has more to gain through waiting chooses a higher adoption intensity. This is a failing of mixed-strategy Nash equilibria in general, and is not peculiar to this model.

© Basil Blackwell Ltd. 1995.

staggered: if the instantaneous adoption intensity is α, the probability of k firms adopting simultaneously in a small interval ε is of order $(\alpha\varepsilon)^k$ which is of lower order than ε for $k > 1$. Once a firm adopts, the others defer adoption by δ in order to gather information, so that firms tend to adopt sequentially rather than simultaneously. Because exactly one firm adopts at each stage, some conjectures about the diffusion path follow directly once we determine the duration of a stage.

Recall that the duration of a stage is given by the time interval between two successive intervals and, in the equilibrium described above, the interval comprises two phases. In the n-th stage-game, the strategic phase lasts from t_n to τ_n, and is followed by an information-gathering phase from τ_n to $\tau_n + \delta$. At this point the next stage begins; that is, we have $t_{n-1} = \tau_n + \delta$. The total duration of the n-th stage can be written as

$$d_n \equiv t_{n-1} - t_n = \delta + (\tau_n - t_n).$$

Since δ is exogenously specified, to determine the duration of this stage we need only to determine the duration $(\tau_n - t_n)$ of the strategic phase. By construction, τ_n is the time of the earliest adoption from among the n firms in that stage-game. Given that we know the distribution over adoption times for each firm, we can obtain the distribution of the earliest adoption time from among the n firms. Since each firm has an exponential distribution over adoption times, it follows from standard properties of the exponential distribution, that the duration up to the earliest adoption is distributed exponentially as well, with parameter equal to the sum of the individual parameters. To formalise this, we define *the aggregate adoption intensity* for a stage-game as the sum of equilibrium adoption intensities of all firms that are currently active. It is written as

$$\Lambda(n, p_n) = \begin{cases} n\hat{\alpha}(n, p_n), & \text{if } \pi_n \geqslant 0 \text{ and } v_t = 0, \\ 0 & \text{otherwise} \end{cases}$$

Note that it would suffice to say that $\Lambda = n\alpha$. The earliest adoption time in stage n is distributed exponentially with parameter Λ. And since the mean of an exponential distribution equals the inverse of its parameter, we have

Proposition 3. The expected duration of the n-th stage is given as

$$\mathrm{E}[d_n] = \delta + [\Lambda(n, p_n)]^{-1}.$$

The next proposition describes how the aggregate adoption intensity varies with n and p_n. We confine our attention to the strategic phase.

Proposition 4. In the strategic phase, the aggregate adoption intensity $\Lambda(n, p_n)$ is increasing in p_n, and decreasing in n.

Propositions 3 and 4 relate the pace of diffusion to the fundamental

© Basil Blackwell Ltd. 1995.

characteristics of the innovation in a manner that is consistent with received wisdom. Ceteris paribus, in terms of this model, the greater is the expected profitability of the innovation, the higher is the aggregate adoption intensity and the shorter is the expected duration of each stage. Mansfield [1961] found substantial evidence that the 'rate of imitation' tended to be higher for profitable innovations. At the same time, the greater the significance of the information that is expected to arrive in the future, the lower is the adoption intensity and therefore, the longer is the expected duration of a stage. In this model, information helps a firm to reduce the risk of adopting 'bad' technologies. The anticipated information is especially significant if the cost of an erroneous adoption is high, so that our model predicts longer diffusion lags for such technologies. Once again, this prediction is consistent with empirical evidence — Mansfield found the rate of imitation to be lower for innovations that required large initial investments.

To show that the adoption lags predicted by this model are not entirely out of line with observed facts, consider the following as a finger exercise. Let the rate of interest be 10% per annum, and suppose, somewhat conservatively, that the proportional advantage from waiting for information, that is $(e^{-\rho\delta}\bar{\pi}_{|.} - \pi)/\pi$, is of the order of 2%. Then, the aggregate adoption intensity is given as

$$\Lambda = \frac{n}{n-1}\frac{0.1}{0.02} \text{ per annum.}$$

The ratio $n/(n-1)$ does not affect the arithmetic in too significant a manner: its value ranges from 2 to 1 as n varies in its permissible range of 2 to any finitely large number. For $n = 2$ the above parameters imply an aggregate adoption intensity of about 10 per annum, and for large n the corresponding figure is of the order of 5 per annum. These values imply a strategic phase with an expected duration between 0.1–0.2 years, or of the order of a few months. Combined with some reasonable specification of the information lag δ, we obtain diffusion lags that are not implausible for many technologies. Note that in the early stages of the diffusion of most technologies, information might be more valuable than suggested by the 2% figure in this exercise. If the figure were higher, the expected duration of the stage would be correspondingly longer.

The implications of this analysis for the pattern of diffusion are fairly transparent. One, it creates the possibility that a run of bad experiences with an innovation, even a good one, might arrest the process of diffusion. This, in turn, would foreclose the generation of fresh information about the technology and potentially good innovations might be lost irreversibly. Two, the conjunction of Propositions 3 and 4 have some bearing on the shape of diffusion curves for good innovations. It is commonly suggested that firms tend to be cautious towards new technologies when they first arrive. One

© Basil Blackwell Ltd. 1995.

possible way to interpret this caution in terms of the model presented here would be to say that at the initial stage N, firms are relatively sceptical of the merits of the technology, or that the value p_N is relatively small. Then, by Proposition 4, the aggregate adoption intensity will be small as well (provided only that p_N is large enough for expected profitability to be positive; otherwise the adoption intensity would be 0). For a good innovation we expect that, with successive adoptions, p will tend towards its true value, namely 1, by the law of large numbers. As p rises, the aggregate adoption intensity tends to increase. (Additionally, the rise in the aggregate adoption intensity is reinforced by the index n falling through time, though this latter effect is relatively small). As a consequence, the expected duration of each stage tends to fall from one stage to the next. In sum, as confidence in the technology grows, we expect the diffusion lags to shrink. Of course, the process of diffusion will be arrested if for any stage in the evolution of beliefs, a run of bad experiences drives down the expected profitability of its adoption to a negative value; this is especially likely when the innovation is a 'bad' one. If we plot the number of adoptions $y_t = N - n_t$ against the (expected) time taken for that number of adoptions to have occurred, the smooth-line curve joining these points could be viewed as an *expected diffusion curve*. If the expected duration of a stage falls over the course of diffusion, successive adoptions follow each other more closely - the curve rises gently at first and then becomes steeper. The informational externality described in this model is, by its very nature, more relevant to the early stages of diffusion. Figure 1 provides a sketch of an expected diffusion curve for a good innovation. Note that this is consistent with the early stages of an S-shaped pattern as described in empirical studies.

Let us now consider the role of the information lag δ on the pace of diffusion which, interestingly, seems to have an ambiguous effect on the pace of diffusion.

Proposition 5. The effect of varying δ on the mean stage-duration $E[d_n]$ is ambiguous.

In intuitive terms, a higher value of δ implies later availability of information after any adoption, and this makes waiting for that information less appealing. This implies a higher adoption intensity in the strategic phase. So we would have a longer information-gathering phase, but the expected duration of the strategic phase would be lower, and the net effect cannot be signed. This may have interesting policy implications. Suppose that, subject to certain limits, the information lag δ is within the ambit of policy control. This may be if, say, the principal channel of information-revelation among firms is published company accounts, and the periodicity of statutory declarations is within the control of the policy maker. A policy maker eager to hasten the diffusion of an innovation that she believes to be good might be tempted

© Basil Blackwell Ltd. 1995.

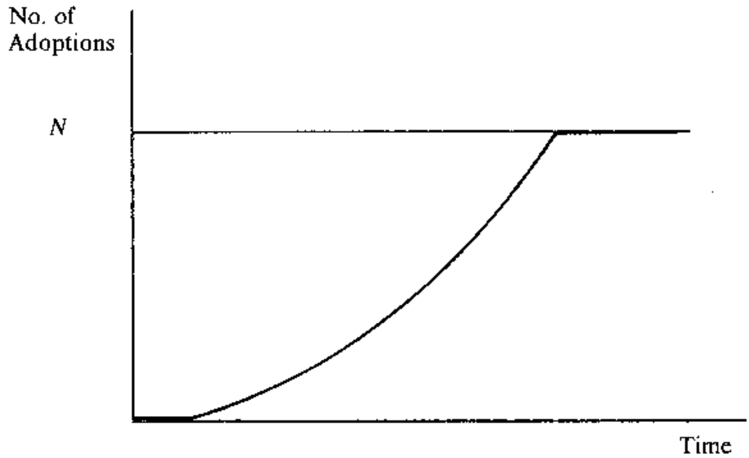

Figure 1
A Diffusion Curve

to lower δ. While this would shorten the information gathering phase, it could prolong the strategic phase. In fact, the strategic phase might be prolonged for sufficiently many stages such that the overall pace of diffusion is lower.

Lastly, note that in this model the delays in the diffusion process are caused by coordination failure. Each firm delays adoption in the hope that, if others were to adopt before they did, it will be able to make a better-informed adoption decision. Some coordination of the firms' intentions, if it could be achieved by some mechanism, would be advantageous. To illustrate this point we consider a simple two-firm example. Suppose an innovation arrives at time $t = 0$ and each firm must choose when to adopt it. The firm that adopts first is designated as the leader and gets π^L; the follower, if it adopts no sooner than δ after the leader, gets $e^{-\rho\delta}\pi^F$; otherwise it gets π^L as well. Assume that $e^{-\rho\delta}\pi^F > \pi^L > 0$, so that each firm would prefer to follow rather than lead. If, as in our model, the adoption timing decision is taken in a decentralised, uncoordinated manner, the equilibrium solution typically involves some lag before *either* firm adopts and each firm's expected gain from the technological opportunity is just π^L in equilibrium. The outcome could be improved by means of 'indicative planning' as described, for instance, in Bolton and Farrell [1990]. It might be the case that a social planner could intervene, randomly nominating a firm to be the leader without any coercive power. This would eliminate the coordination problem, and the consequent delay, so that one firm would adopt immediately, and the other would follow with a lag δ. If each firm has equal probability of being nominated the leader, the expected payoff for each firm would be $0.5(e^{-\rho\delta}\pi^F + \pi^L)$ which, given our assumption, is larger than π^L. The same outcome might obtain if the

© Basil Blackwell Ltd. 1995.

government, as the social planner, used its general power of taxation to provide a *selective* subsidy to a pre-nominated firm. Generalisations of this argument to the N-firm cases are less than straightforward since the optimal strategy for the planner is likely to be sensitive to the speed of learning, so that clear policy prescription must await further research.

IV. CONCLUSIONS

To summarise, while a large body of the literature on diffusion attributes inter-firm differences in adoption times to the inherent heterogeneity of firms, this paper argues that heterogeneity is not essential for diffusion. Of course, the claim is not unique to this model. Other game-theoretic models demonstrate how identical firms could end up adopting at different times but their analyses remain limited when it comes to the determinants of the pace of diffusion. This paper remedies the lacuna in some measure. The central focus is on the informational advantage to delayed adoption: the ability to observe others' success or failure with an innovation confers an advantage on those who adopt after others, in as much as they can reconsider their decision in the light of the new information. The process of diffusion is modelled as a sequence of 'wars of attrition' among the potential adopters at any stage, and these stage-games are linked together by a Bayesian updating of beliefs. We show that, in equilibrium, each firm employs a mixed strategy over adoption times, and the intensity of adoption depends on the perceived characteristics of the technology and on the learning process. The mean duration of each stage is determined endogenously, allowing us to construct some crude diffusion paths.

Our model concentrates exclusively on an informational externality in the adoption process. In particular, the analysis assumes that the private gain from adoption is independent of the technological choices of other firms. Clearly, this is restrictive. For instance, early entrants often capture the best location in a geographical setting; pioneers usually acquire brand-recognition as a technological leader. Would the results be altered drastically if we take account of such issues? The spirit of the analysis remains valid as long as the informational advantage to delayed adoption dominates the possible advantages to early adoption. In general, Condition 1 is less likely to be valid, but if it were valid, the logic of Proposition 1 carries through. What changes is the manner in which the stage-games tie together. A little reflection suggests that the shape of the expected diffusion curve depends on the progressive variation in the gain to deferring adoption *relative* to that from immediate adoption. At the risk of seeming to rely on ad hoc assumptions, we can support a wide variety of diffusion patterns. For instance, suppose that as the innovation diffuses, the increasing experience with it not only reduces the initial uncertainty but also creates the possibilty of future improvements in it. This may happen if the accumulated experience with the

© Basil Blackwell Ltd. 1995.

innovation induces the technology suppliers to introduce an improved version of the innovation. If the likelihood of an improved version arriving is increasing in the level of accumulated experience, this tendency could raise the expected gain from waiting in the later stages of diffusion. This might even support the S-shaped diffusion curve within our model. To see this, consider the following possibility as a thought experiment. In the early stages, the pace of diffusion could be influenced primarily by the possibility of learning more about the technology. As the uncertainty diminishes, the potential gain from waiting tends to decline, causing the pace of diffusion to rise. However, after some stage, the expected gain from waiting could increase once again, principally because of the growing anticipation that an improved technology will soon become available. This anticipation would then tend to retard the pace of diffusion, causing the diffusion curve to taper off.

In this model the delays in the process of diffusion are, to some extent, induced by coordination failure. To clarify the issue, it must be pointed out that there are two sources of uncertainty that affect the firms' decisions in this model. First, there is the uncertainty about the nature of the technology. This technological uncertainty is exogenous to the firms' decisions and is faced by the industry as a whole. Additionally, from the viewpoint of each firm, there is the uncertainty regarding the timing of its rivals' adoptions. This sort of uncertainty is endogenous to the diffusion process and (for the want of a better term) may be called *strategic uncertainty*. The sequential nature of adoption is optimal in relation to the technological uncertainty — it allows firms to learn from each others' mistakes — but the strategic uncertainty results in adoption lags that are 'excessively' long. Put simply, the diffusion path tends to be more time-intensive than would be strictly necessary for learning in our model. In such situations, some gains could possibly result from the better co-ordination of firms' information and intentions.

SANDEEP KAPUR,
Department of Economics,
Birkbeck College,
University of London,
Gresse Street,
London W1P 1PA,
UK

ACCEPTED NOVEMBER 1994

APPENDIX

Proposition 1. Given Condition 1, as $\Delta \to 0$, the symmetric Nash equilibrium of $G_\Delta(n, p_n)$ is given as $[\alpha(n, p_n), \alpha(n, p_n), \ldots, \alpha(n, p_n)]$, such that

$$\alpha = \begin{cases} \hat{\alpha}(n, p_n) & \text{if } \pi_n \geqslant 0 \text{ and } v_t = 0; \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{where } \hat{\alpha}(n, p_n) \equiv \frac{1}{(n-1)} \frac{\rho \pi_n}{(e^{-\rho\delta} \bar{\pi}_{|n,1} - \pi_n)}.$$

© Basil Blackwell Ltd. 1995.

Proof: First consider the case where p_n is such that $\pi_n < 0$. Clearly, it is not optimal for any firm to ever adopt at such a stage: adoption yields a negative expected return, while waiting for information guarantees at least a nil return. Hence, each firm chooses $\beta^* = 0$ for all (t, v_t, n). Since this value is independent of Δ, in the limit too we have $\alpha = \text{Lim}_{\Delta \to 0} \beta^*/\Delta = 0$. Intuitively, pessimism about the innovation halts the process of diffusion altogether.

Next consider the case where $\pi_n \geqslant 0$ and $v_t \neq 0$. The former implies that adoption is profitable, but the latter indicates that at least one firm adopted the technology in the interval $[t - m\Delta, t - \Delta]$. A signal is surely expected before $t + \delta$ allowing the firm to take an informed decision by $t + \delta + \Delta$: given condition 1 it must be optimal to wait for that signal. Once again, we have $\beta^* = 0$, and consquently $\alpha = 0$.

Lastly, consider the case where $\pi_n \geqslant 0$ and $v_t = 0$ so that no information is expected in the interval $(t, t + \delta)$. Notice that as long as no firm adopts the game is stationary — if t is reached, the game looks the same as it was at $t - \Delta$. We consider symmetric equilibria of this game in behaviourally-mixed strategies. At such an equilibrium, each firm is indifferent between adoption and waiting. The expected gain from adoption equals π_n, while that from waiting depends on the actions chosen by the rival firms. Let β be the (symmetric and stationary) probability that a typical rival will adopt in an aribtrary time period. The gain from waiting at any time period can be written as

$$W_n = (1 - \beta)^{n-1} e^{-\rho\Delta} \phi_n(p_n) + \sum_{k=1}^{n-1} {}^{n-1}C_k \beta^k (1 - \beta)^{n-1-k} e^{-\rho(\delta+\Delta)} \phi_{n-k}(p_n).$$

The right hand side in this expression is the probability-weighted sum of the value of ending up in various continuation stage-games. To understand this, note that if none of the other $n - 1$ firms adopts at t, the firm will find itself in the (unchanged) n-th stage-game. Its value $\phi_n(p_n)$ is discounted for the lag Δ and weighted with the probability $(1 - \beta)^{n-1}$ of that event to obtain the first term. On the other hand, if $1 \leqslant k \leqslant n - 1$ firms adopt at t, we have already determined that the equilibrium strategy involves waiting for the resulting signal to arrive. Once it arrives the firm finds itself in the continuation game corresponding to stage $n - k$ and the value of being in that stage-game depends on the posterior beliefs. Let $\phi_{n-k}(p_n)$ be the expected value of being in the $(n - k)$th stage-game, the expectation being over all possible transitions of p_n. Discounting this value for the information lag δ and the reaction lag Δ, and weighting it with the probability of k adoptions, we get the typical term in the summation.

If a mixed strategy equilibrium exists, at that equilibrium each firm must be indifferent between waiting and adoption. We have

$$W_n = \phi_n = \pi_n.$$

Using this in the previous expression, we have

$$\pi_n = (1 - \beta)^{n-1} e^{-\rho\Delta} \pi_n + \sum_{k=1}^{n-1} {}^{n-1}C_k \beta^k (1 - \beta)^{n-1-k} e^{-\rho(\delta+\Delta)} \phi_{n-k}(p_n),$$

or multiplying both sides by $e^{\rho\Delta}$, we get

$$\text{(A1)} \qquad e^{\rho\Delta} \pi_n = (1 - \beta)^{n-1} \pi_n + \sum_{k=1}^{n-1} {}^{n-1}C_k \beta^k (1 - \beta)^{n-1-k} e^{-\rho\delta} \phi_{n-k}(p_n).$$

The proof is now in three steps. First, we show that given condition 1 and $\Delta > 0$, there exists a β in the open interval $(0, 1)$ for which (A1) holds. Then it is argued that, in equilibrium ϕ_{n-k}, the expected value of being in the continuation game, equals $\bar{\pi}_{|n,k}$. This allows us to argue that the solution to (A1) in the unit interval (call it β^*_Δ)

© Basil Blackwell Ltd. 1995.

is unique. Next, as $\Delta \to 0$ equation (A1), which is of degree $n-1$ in β, is approximated by a linear equation, the solution to which yields the equilibrium adoption intensity $\hat{\alpha}$.

Lemma 1: For $k = n-1$, we have $\phi_{n-k} = \bar{\pi}_{|n,k}$. That is, if $n-1$ firms adopt simultaneously, the expected value of being in the subsequent stage-game is $\bar{\pi}_{|n,n-1}$.

Proof: If all firms except the reference firm adopt simultaneously, that firm faces a single-agent optimisation problem, and the best it can do is to react optimally to the information provided by those firms. By definition, the expected value of that is $\bar{\pi}_{|n,n-1}$. □

Lemma 2: Equation (A1) has a solution in the unit interval for all n.

Proof: Rewrite (A1) as follows:

$$\text{(A2)} \qquad [(1-\beta)^{n-1} - e^{\rho\Delta}]\pi_n + \sum_{k=1}^{n-1} {}^{n-1}C_k \beta^k (1-\beta)^{n-1-k} e^{-\rho\delta} \phi_{n-k}(p_n) = 0.$$

This is of the form $g(\beta;\Delta) = 0$, where g is a polynomial function of degree $n-1$ in β, and g is parametrised by Δ. We have

$$g(0;\Delta) = [1 - e^{\rho\Delta}]\pi_n, \qquad \text{and using Lemma 1 we have}$$

$$g(1;\Delta) = -e^{\rho\Delta}\pi_n + e^{-\rho\delta}\bar{\pi}_{|n,n-1}.$$

Clearly, for $\Delta > 0$, we have $g(0;\Delta) < 0$, and given condition 1, $g(1;\Delta) > 0$. Since g is a polynomial function, and therefore continuous, there must exist some $\beta^*_\Delta \in (0,1)$ such that $g(\beta^*_\Delta) = 0$. □

Lemma 3: $\phi_{n-k}(p_n) = \bar{\pi}_{|n,k}$ for all k.

Proof: Note that Lemma 2 holds for any arbitrary n. Suppose k forms adopt simultaneously so that, given $s \in S_k$, the beliefs for the subsequent $(n-k)$th stage-game are given by $\tilde{p}$. The equilibrium strategy in that stage-game dictates (i) not adopt if $\pi(\tilde{p}) < 0$, which yields a zero payoff; or (ii) a non-trivial randomisation over the adoption decision if $\pi(\tilde{p}) \geqslant 0$, and if condition 1 holds. The latter must provide the same expected return as playing any pure strategy in the support (e.g., adoption) which yields $\pi(\tilde{p})$. Taking expectations over S_k, and using the definition of $\bar{\pi}_{|n,k}$, the claim follows. □

Given Lemma 3 we can rewrite (A2) as

$$\text{(A3)} \qquad [(1-\beta)^{n-1} - e^{\rho\Delta}]\pi_n + \sum_{k=1}^{n-1} {}^{n-1}C_k \beta^k (1-\beta)^{n-1-k} e^{-\rho\delta} \bar{\pi}_{|n,k} = 0.$$

Lemma 4: The solution to (A3) in the unit interval, β^*_Δ, is unique.

Proof: Differentiate $g(\beta)$, as in the left hand side of (A3) with respect to β. Collecting terms we get

$$\frac{\partial g}{\partial \beta} = $$

$$(n-1)\left[(1-\beta)^{n-2}(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n) + \sum_{k=1}^{n-2} {}^{n-2}C_k \beta^k (1-\beta)^{n-2-k} e^{-\rho\delta}(\bar{\pi}_{|n,k-1} - \bar{\pi}_{|n,k})\right].$$

© Basil Blackwell Ltd. 1995.

From Condition 1 we know that $(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n) > 0$. And given that $\bar{\pi}_{|n,k}$ is non-decreasing in k — recall that higher k implies a finer information structure so the value of conditioning on that information cannot be lower — we have $(\bar{\pi}_{|n,k+1} - \bar{\pi}_{|n,k}) \geqslant 0$ for all k. It follows that the derivative must be positive in the interval $[0, 1)$; in particular it equals $(n-1)(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n)$ at $\beta = 0$. Because $g(\beta; \Delta)$ is strictly monotonic in the unit interval, the solution to $g(\beta; \Delta) = 0$ in the unit interval must be unique. □

We now consider what happens to this solution as the decision interval shrinks. We are interested in the instantaneous intensity $\text{Lim}_{\Delta\to 0}\, \beta^*_\Delta/\Delta$. To evaluate this, note that $g(\beta; \Delta) = 0$ is of the form

(A4) $$K_{n-1}\beta^{n-1} + \ldots + K_1\beta + K_0(\Delta) = 0$$

where $K_1 = (n-1)(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n)$, and the only term that varies with Δ is $K_0 = (1 - e^{\rho\Delta})\pi_n$. For $\Delta = 0$, we have $K_0 = 0$ so that (A4) reduces to

$$K_{n-1}\beta^{n-1} + \ldots + K_1\beta = 0,$$

with root $\beta^* = 0$ in the unit interval. That implies that β^*_Δ/Δ evaluated at $\Delta = 0$ is of the indeterminate form 0/0. By L'Hopital's rule

$$\text{Lim}_{\Delta\to 0} \frac{\beta^*_\Delta}{\Delta} = \text{Lim}_{\Delta\to 0} \frac{\partial\beta^*_\Delta/\partial\Delta}{\partial\Delta/\partial\Delta}.$$

We do not have an explicit expression for β^*_Δ; however, by the implicit differentiation of $g(\beta; \Delta) = 0$, we know $\partial\beta^*_\Delta/\partial\Delta = -g_\Delta/g_\beta$. Now, recalling that K_0 is the only term that varies with Δ so that $g_\Delta = \partial K_0/\partial\Delta$. And g_β, as evaluated in the proof of Lemma 4, is positive. Hence $-g_\Delta/g_\beta$ is well defined and finite. We can divide both sides of (A4) by Δ and rewrite as

$$K_{n-1}\left(\frac{\beta}{\Delta}\right)^{n-1}\Delta^{n-2} + \ldots + K_2\left(\frac{\beta}{\Delta}\right)^2\Delta + K_1\left(\frac{\beta}{\Delta}\right) + \frac{K_0}{\Delta} = 0.$$

Note that as $\Delta \to 0$, all higher order terms vanish and the equation reduces to

$$K_1\left(\frac{\beta}{\Delta}\right) + \frac{K_0}{\Delta} = 0, \qquad \text{so that } \frac{\beta^*_\Delta}{\Delta} = \frac{-K_0(\Delta)}{\Delta K_1} = \frac{-(1 - e^{\rho\Delta}\pi_n)}{\Delta(n-1)(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n)}.$$

Taking the limits, once again by L'Hopital's rule, we have

$$\hat{\alpha} = \text{Lim}_{\Delta\to 0} \frac{\beta^*_\Delta}{\Delta} = \frac{1}{(n-1)} \frac{\rho\pi_n}{e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n} \qquad \text{as stated.}$$ □

Proposition 2. $\hat{\alpha}(n, p)$ is increasing in p.

Proof: We write $\pi_n = \pi(p)$ and $\bar{\pi}_{|n,1} = \bar{\pi}(p)$, so that

$$\hat{\alpha}(n, p) = \frac{1}{(n-1)} \frac{\rho\pi(p)}{[e^{-\rho\delta}\bar{\pi}(p) - \pi(p)]}.$$

Both the numerator $\pi(p)$ and the denominator are positive, and since $\pi(p) = p\theta_h + (1-p)\theta_l$ is increasing in p, it is sufficient to show that the denominator is non-increasing in p. That is, it is sufficient to show that

$$[e^{-\rho\delta}\bar{\pi}(p') - \pi(p')] \leqslant [e^{-\rho\delta}\bar{\pi}(p'') - \pi(p'')] \qquad \text{for } p' > p''.$$

In fact since $e^{-\rho\delta} < 1$, it is sufficient to show that

© Basil Blackwell Ltd. 1995.

$$[\bar{\pi}(p') - \pi(p')] \leqslant [\bar{\pi}(p'') - \pi(p'')] \quad \text{for} \quad p' > p''.$$

To do that, we need a further construct. Define p_c such that $\pi(p_c) = 0$; the value p_c defines the break-even level of confidence in the technology. For any p we define the following partition on the set of signals S,

$$\bar{S}(p) = \{s \in S : \tilde{p}(s, p) \geqslant p_c\};$$

$$\underline{S}(p) = \{s \in S : \tilde{p}(s, p) < p_c\}.$$

For a given p, $\bar{S}$ is the set of signals which, if received would lead to the posterior belief $\tilde{p}$ being at least as large as p_c, and $\underline{S}$ is the set of signals for which the posterior is less than p_c. Now, by definition

$$\bar{\pi}(p) = \sum_{s \in S} \Pr(s) \operatorname{Max}[\pi(\tilde{p}(s, p)), 0], \text{ which, using the above partition of } S$$

$$= \sum_{s \in \bar{S}(p)} \Pr(s)\pi(\tilde{p}(s, p)).$$

Using Bayes Law to obtain expressions for $\Pr(s)$ and $\tilde{p}$, we get

$$\bar{\pi}(p) = \sum_{s \in \bar{S}(p)} [p\sigma_h(s)\theta_h + (1 - p)\sigma_l(s)\theta_l].$$

In contrast,

$$\pi(p) = \sum_{s \in S} [p\sigma_h(s)\theta_h + (1 - p)\sigma_l(s)\theta_l],$$

where the difference in this and the previous expression lies in that the latter summation is carried out over S rather than $\bar{S}$. This implies

$$\bar{\pi}(p) - \pi(p) = \sum_{s \in \underline{S}(p)} - [p\sigma_h(s)\theta_h + (1 - p)\sigma_l(s)\theta_l] \equiv Z(p), \text{ say.}$$

Now consider any p', p'' such that $p' > p''$. By construction, we must have $\underline{S}(p') \subseteq \underline{S}(p'')$.

$$Z(p'') - Z(p') = \sum_{s \in \underline{S}(p')} (p' - p'')[\sigma_h(s)\theta_h - \sigma_l(s)\theta_l]$$
$$- \sum_{s \in \underline{S}(p'')/\underline{S}(p')} [p''\sigma_h(s)\theta_h + (1 - p'')\sigma_l(s)\theta_l].$$

The first term is positive since $p' > p''$ and $\theta_l < 0 < \theta_h$; the term inside the second summation is negative since $s \in \underline{S}(p'')$. Therefore, $Z(p'') > Z(p')$, as required. □

Proposition 3. The expected duration of the n-th stage $\mathrm{E}[d_n] = \delta + [\Lambda(n, p_n)]^{-1}$.

Proof: The information-gathering phase lasts δ. The expected duration of the strategic phase equals Λ^{-1}: this follows from the standard properties of the exponential distribution described, for instance, in Mood, Greybill, and Boes [1974].

Proposition 4. In the strategic phase, the aggregate adoption intensity $\Lambda(n, p_n)$, is increasing in p_n and decreasing in n.

Proof: By definition, in the strategic phase

$$\Lambda(n, p_n) = \frac{n}{(n - 1)} \frac{\rho\pi(p)}{[e^{-\rho\delta}\bar{\pi}(p) - \pi(p)]}.$$

The first part of the claim follows directly from Proposition 2, and second part form the fact that $n/(n - 1)$ is decreasing in n. □

© Basil Blackwell Ltd. 1995.

Proposition 5. The effect of varying δ on the mean stage-duration $E[d_n]$ is ambiguous.

Proof: Differentiating $E[d_n]$ with respect to δ, we get

$$\frac{\partial}{\partial\delta}E[d_n] = 1 - \frac{(n-1)}{n}\frac{(e^{-\rho\delta}\bar{\pi}_{|n,1})}{\pi_n} > 0 \text{ if and only if } n < \frac{e^{-\rho\delta}\bar{\pi}_{|n,1}}{(e^{-\rho\delta}\bar{\pi}_{|n,1} - \pi_n)}. \quad \square$$

REFERENCES

BALCER, YVES and LIPPMAN, STEVEN, 1984, 'Technological Expectations and the Adoption of Improved Technology', *Journal of Economic Theory*, 34, pp. 292–318.

BESLEY, T. and CASE, A., 1993, 'Modeling Technology Adoption in Developing Countries', *American Economic Review, Papers and Proceedings*, 1993, pp. 396–402.

BHATTACHARYA, S., CHATTERJEE, K. and SAMUELSON, L., 1986, 'Sequential Research and the Adoption of Innovations', *Oxford Economic Papers*, 1986, pp. 219–243.

BLACKWELL, D., 1951, 'Comparison of Experiments', in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press).

BOLTON, PATRICK and FARRELL, JOSEPH, 1990, 'Decentralization, Duplication, and Delay', *Journal of Political Economy*, 98, no. 4, pp. 803–826.

CHAMLEY, C. and GALE, D., 1994, 'Information Revelation and Strategic Delay in a Model of Investment', *Econometrica*, 62, No. 5, pp. 1065–1085.

DASGUPTA, PARTHA, 1986, 'The Theory of Technological Competition', in BINMORE, K. and DASGUPTA, P. (eds), *Economic Organizations as Games* (Basil Blackwell, Oxford).

DASGUPTA, PARTHA, 1988, 'Patents, Priority and Imitation or, The Economics of Races and Waiting Games', *The Economic Journal*, 98, pp. 66–80.

DAVIES, S., 1979, *The Diffusion of Process Innovations*, (Cambridge University Press, Cambridge).

ELLISON, G. and FUDENBERG, D., 1993, 'Rules of Thumb for Social Learning', *Journal of Political Economy*, vol. 101, no. 4, pp. 612–643.

FUDENBERG, D. and TIROLE, J., 1985, 'Pre-emption and Rent Equalization in the Adoption of New Technology', *Review of Economic Studies*, LII, pp. 383–401.

FUDENBERG, D., and TIROLE, J., 1991, *Game Theory*, (MIT Press, London).

GRILICHES, ZVI, 1957, 'Hybrid Corn: An Exploration in the Economics of Technological Change', *Econometrica*, 25, pp. 501–22.

JENSEN, R., 1982, 'Adoption and Diffusion of an Innovation of Uncertain Profitability', *Journal of Economic Theory*, 27, pp. 182–193.

KAPUR, S., 1994, 'Markov-Perfect Equilibria in a N-player War of Attrition', *Economics Letters*, forthcoming.

MANSFIELD, E., 1961, 'Technical Change and the Rate of Imitation', *Econometrica*, pp. 741–66.

MANSFIELD, E., 1968, *The Economics of Technical Innovation*, (WW Norton, New York).

MARIOTTI, MARCO, 1992, 'Unused Innovations', *Economics Letters*, 38, pp. 367–371.

MAYNARD SMITH, J., 1974, 'The Theory of Games and the Evolution of Animal Conflicts', *Journal of Theoretical Biology*, 47, pp. 209–221.

MOOD, A., GREYBILL, F. and BOES, D., 1974, *Introduction to the Theory of Statistics*, Third edition, (McGraw-Hill).

REINGANUM, J. F., 1989, 'The Timing of Innovation: Research, Development and Diffusion', in SCHMALENSEE, R. and WILLIG, R. D. (eds.), *Handbook of Industrial Organization* (North-Holland, Amsterdam).

© Basil Blackwell Ltd. 1995.

REINGANUM, J. F., 1985, 'A Two Stage Model of Research and Development with Endogenous Second-mover Advantages', *International Journal of Industrial Organization*, 3, pp. 275–292.

ROSENBERG, N., 1976, 'On Technological Expectations', *Economic Journal*, 86, pp. 523–535.

© Basil Blackwell Ltd. 1995.