

NBER WORKING PAPER SERIES

COMPETITIVE DIFFUSION

Boyan Jovanovic
Glenn MacDonald

Working Paper No. 4463

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1993

This re-h was supported by the National Science Foundation, the Lynde and Harry Bradley Foundation, the John M. Olin Foundation, and the CV. Starr Center for Applied **Economics**. Phil Reny, Sherwin Rosen, and Yaw Nyarko provided helpful comments on an earlier draft. **This** paper is part of **NBER's** research program in Productivity. Any opinions expressed are **those** of the authors and not those of the National Bureau of Economic **Research**.

COMPETITIVE DIFFUSION

ABSTRACT

The usual explanation for why the **producers** of a **given** product **use** different technologies involves "vintage-capital": A firm understands the frontier technology, but can still prefer an older, less efficient technology in which it has made specific physical and human capital investments. This paper develops an alternative. "information-barrier" hypothesis: Firms differ in the technologies they use because it is costly for them to overcome the informational barriers that separate them. The paper **endogenizes** both innovative and imitative effort. The industry life-cycle implications -- declining price and increasing output -- broadly agree with the **Gort-Klepper** data.

Empirically, the **paper** focuses on the slow spread of Diesel locomotives, which can not be explained by the vintage-capital hypothesis alone. For instance, contrary to **that** hypothesis, railroads were buying new steam locomotives long after the Diesel first came into use -- exactly as the **information-barrier** hypothesis would imply.

Boyan Jovanovic
Department of Economics
New York University
New York, NY 10003
and NBER

Glenn MacDonald
Simon School of Business
University of Rochester
Rochester, NY 14627

Solow (1957) observed that most growth in economic activity cannot be **explained** by increasing quantities of inputs producing output with a **fixed technology**. This paper models the unexplained component and **proposes** a theory of the development and spread of new **technology** at the industry level. The theory is motivated by evidence like Figure 1. Panel a) **displays** the fraction of shipments of bits of dynamic random access memory by chip density and over time. Low density (1 kilobyte) chips are displaced by those with higher density (4 k), which are then taken over by those with yet higher density (16, then 64 and then 256 k). Meantime, the quantity of bits delivered explodes and price declines dramatically. The bit industry displays the waves of change and improvement stressed by **Schumpeter** (1934). Such data call for a theory in which new developments occur periodically and do not spread instantaneously.

In the setup studied here, the growth and diffusion of technology are **both** endogenous. Firms improve their know-how both by producing new knowledge-- innovation-- and by learning from others-- imitation. Both activities are costly and thus respond to the incentives provided by the economic environment.¹

Informational barriers appear to play an important role in explaining lags in the adoption of technology: **Nabseth** and Ray (1974) and Rogers (1983) report that some **firms** learn of a relevant new technology more than a decade before others. Moreover, being aware of a technology is not the same as mastering it-- According to Mansfield, Schwartz and Wagner (1981) and Pakes and **Shankerman** (1984), imitation and product-development lags are long. Thus, the spread of information seems to be gradual and costly. The hypotheses developed in this paper flow from a learning process embodying informational barriers and implying costly and gradual learning in equilibrium along with a non-degenerate distribution of technological knowledge among firms in an industry.²

¹ The list of attempts to **endogenize** technological progress is long. At the macro level, see **Arrow** (1962), **Shell** (1967), **Lucas** (1988) and **Romer** (1990); at the micro level, **Flaherty** (1980) and **Spence** (1984).

² **Jovanovic** and **Rob** (1989) also use informational barriers to generate a persistent non-degeneracy in the distribution of technological knowledge, but they do not allow agents to

The vintage capital model is the primary alternative explanation for why firms use different technologies; see, for example, **Chari** and Hopenhayn (1991). In that model, **firms** have complete technological know-how, but use less than state-of-the-art technology when doing so is profitable given the existing **stock** of assets specific to older technology. Here, **firms** use different technologies because it is costly to overcome the informational barriers that define firm boundaries, Innovation and imitation are alternative ways to make progress, and the relative desirability of each depends on the firm's current know-how and the know-how of others.

The paper contains three types of results. The **first** three propositions deal with the evolution of industry aggregates, and describe the precise sense in which technology improves, output increases, and price declines as the **industry** matures. The technological diversity of **firms** may or may not persist forever, depending on the learning technology.

The second set of results focuses on innovative behavior in a cross section of firms at a point in time. Under some conditions, small firms will, on average, grow faster than big firms. There are two reasons why laggards may grow faster than leaders. The **first** is the diffusion of technology from leaders to followers via imitation. The second is a cross-sectional "fishing out effect": if all **firms** are sampling new technologies from the same pool, leaders, who have already acquired better technology, have less incentive to look for even better ones. As a result, laggards look harder, and this causes at least partial convergence of output and technology over **firms** as the industry matures.

The third set of results deals with the **optimality** of equilibrium. Imitation creates technological spillovers which are akin to an externality, so equilibrium is not generally "efficient". While a global optimum is not analyzed fully, there is a sense in which a social planner would prefer more of all learning activities, **both** innovative and imitative.

These results are first presented in a general form. Then, a series of examples serve to illustrate various aspects of the general results, and to point to some conclusions that do not

substitute between innovation and imitation.

emerge at the **general level**. **One** of these results is that **innovation** and imitation tend to be *substitutes* -- a firm relies mainly on one **or** the other form of cost-reduction -- **and given the** way R&D spending is measured, this substitution possibility is **likely** to bias downward the estimated rate of return to R&D. Another result is that diffusion-- an "equalizer"-- is triggered **by** technological inequality **which** spawns the imitative effort **needed to** drive the diffusion.

All these results hold in a partial equilibrium environment without entry or exit. **Jovanovic** and MacDonald (1993) simplify this model, but also **allow** for entry **and exit and** estimate the resulting model on data from the U.S. automobile tire industry. Andolfatto and MacDonald (1993) simplify too, but then embed the model in a general equilibrium framework. Endogenous growth and cycles emerge; the model is estimated on aggregate U.S. post-war data.

I. Model

The model describes the evolution of a competitive market for a homogeneous product. Demand is subject to exogenous random shocks as a result of income growth, variation in the prices of related goods, etc. Supply is affected by random shocks too, but its development is also governed by firms' introducing cost reducing technological improvements. Firms may get better techniques either directly, through R & D, or indirectly by adopting methods already in use by others. Both of these activities are costly and do not have fully predictable consequences.

Assumptions

Assume discrete time and an infinite horizon. During each period a homogeneous product is sold in a competitive market. Demand for the product is given by an *inverse* industry demand function $D(Q, \mathbf{x}_t)$, where Q , is the quantity produced at date t and \mathbf{x}_t is a **vector** of demand shocks. D is downward sloping and continuous in Q , and \mathbf{x}_t is a realization of a **Markov** process X , which has transition law $F(\mathbf{x}_{t+1} | \mathbf{x}_t)$; \mathbf{x}_0 is given.

In addition to the **homogeneous** product interpretation emphasized here, Q can also be thought of as a flow of services, and technologies as alternative ways to provide consumers **with**

the services they ultimately desire. Thus, the model can accommodate product innovation as well as process innovation.

The supply side of the model comprises a fixed **continuum** of price taking firms who maximize expected discounted profits. To do so they choose a rate of output along with cost reduction effort. There is no entry or exit.

Let a **firm's** state of technological knowledge be denoted by θ , where $0 \leq \theta \leq 1$, and let the cross-firm distribution of know-how at t be ν_t , with ν_0 given.³ Thus, $\nu_t(\theta')$ is the fraction of **firms** in the industry at date t having $\theta \leq \theta'$.

The **firm's** actions will be represented by a vector (q, l) , where q is its rate of output and l is its learning efforts, including the level of R & D in the usual sense, efforts to evaluate others' products, as well as less obvious endeavors such as experimentation with alternative ways to compensate employees or structure financial arrangements with suppliers or distributors. At date t , the **firms'** net revenue is

$$p_t q_t - c(q_t, l_t, \theta, x_t). \quad (1)$$

Here x represents input prices, patent laws, the cost of researchers and of other products to be analyzed, etc. There are no **fixed** costs (i.e. $c(0, 0, \theta, x) = 0$, where $0 = (0, \dots, 0)$), and costs are strictly increasing and convex in (q, l) , and strictly decreasing in θ when $q \neq 0$. A larger θ

³ At the outset, each **firm's** knowledge θ_0 , and hence the distribution of knowledge over **firms**, $\nu_0(\theta)$, is given **exogenously**. The nature of the distribution ν_0 will depend on how new the technology is. By assumption, the **product** is new, but the technology used to produce it need not be entirely new.

If the technology is entirely new, then the initial distribution of knowledge may be concentrated at one value of θ -- "the primitive technology". If, on the other hand, the technology is related to other technologies used to produce older products, then there may be dispersion in ν_0 . But even if firms are technologically all alike at date zero, they will soon become different because the outcomes of **firms'** innovation efforts are random and imperfectly correlated.

therefore denotes a better production technique. Further, $\partial c/\partial q$ is declining in θ ; in this case, given p , better production techniques imply greater Output, and firms can be described as “larger” or “more technologically advanced” interchangeably.’

To specify the evolution of θ , let $\Psi(\theta' | q, l, \theta, x, \nu)$ be the probability that the firm’s know-how next period is less than or equal to θ' . Thus, learning depends on the firm’s state and actions, and on the state of the industry, including the distribution of know-how in use.

The learning technology Ψ satisfies four conditions: (i) the firm cannot guarantee that it will learn ($\Psi(\theta | q, l, \theta, x, \nu) > 0$); (ii), there is no “free” learning ($\Psi(\theta | 0, 0, \theta, x, \nu) = 1$); (iii) the firm’s know-how does not deteriorate ($\Psi(\theta' | q, l, \theta, x, \nu) = 0$ if $\theta' < \theta$); and (iv) increases in the firm’s efforts or know-how, as well as improvements in the know-how of others, add to learning possibilities (if $\hat{q} \geq q$, $\hat{l} \geq l$, $\hat{\theta} \geq \theta$, and $\hat{\nu}$ dominates ν , then $\Psi(\theta' | \hat{q}, \hat{l}, \hat{\theta}, x, \hat{\nu})$ dominates $\Psi(\theta' | q, l, \theta, x, \nu)$.⁵

Maximization

Suppose that the state of the industry can be summarized by the pair (x_t, ν_t) , and that there exist the following equilibrium relationships at each t : $p_t = P(x_t, \nu_t)$ and $\nu_{t+1} = \Phi(x_t, \nu_t)$. Then, given equation (1) and the evolution of the state vector, the firm’s optimal actions solve the dynamic programming problem summarized by

⁴ That technology can be ranked by a scalar is a nontrivial restriction. Some technologies cannot be ordered this way: A labor-saving technique may be superior when wages are high, but need not be better when wages are low. Allowing for this in the present setup would entail that the ranking of technology depend on x as well as θ .

⁵ Whenever distribution functions are being compared, “dominance” will mean first order stochastic dominance, and “improving” will refer to increasing in the sense of first order stochastic dominance.

$$V(\theta, x, \nu) = \max_{(q, l)} \left\{ P(x, \nu)q - c(q, l, \theta, x) + \beta \int V[\theta', x', \Phi(x, \nu)] d\Psi(\theta' | q, l, \theta, x, \nu) dF(x' | x) \right\}, \quad (2)$$

where β is the discount factor and $V(\theta, x, \nu)$ the value of the firm. The first two terms on the right-hand side are the firm's cut-rent profit, while the integral is the expected present value one period later, noting that next period's θ and x are not known at present, and that they may change. Let $(q(\theta, x, \nu), l(\theta, x, \nu))$ be the (unique) action that achieves the maximum in the right hand side of (2); (q, l) is the firm's "policy function".

Equilibrium

Equilibrium demands that firms select optimal policies, that supply equal demand, and that expectations about the evolution of know-how are confirmed: A stationary equilibrium is a set of functions $\{V, q, l, P, \Phi\}$ such that (i) V solves (2); (ii) (q, l) is the policy associated with this solution; (iii) $P(x, \nu) = D[\int q(\theta, x, \nu) d\nu(\theta), x]$; and $\Phi(x, \nu) = \int \Psi[\cdot | q(\theta, x, \nu), l(\theta, x, \nu), \theta, x, \nu] d\nu(\theta)$.⁶

Implications

The analysis of implications is greatly simplified by ignoring learning-by-doing. Learning-by-doing complicates because, as the industry evolves, price tends to fall, giving firms an incentive to shrink. With learning-by-doing, the tendency to contract output implies a tendency to slow the learning process; the analysis becomes more complex and less intuitive.

Thus, assume $c(q, l, \theta, x) = c^q(q, \theta, x) + c^l(l, x)$ and that Ψ is independent of q .

⁶ Jovanovic and MacDonald (1991) prove existence and uniqueness of the policy function and of equilibrium in this environment. and provides a collection of results on the long run behavior of the industry. The technical assumptions made to facilitate that analysis-- for example, regularity conditions used to prove continuity of optimal policies- are suppressed here.

The following general result underlies the time-series implications set out **below**.⁷ It states that aggregate knowledge advances gradually, and never stops entirely:

PROPOSITION 1: Assume **learning begins** at some point; i.e., $\nu_t \neq \nu_0$ for some t . Then for any **sequence** $\{x_t\}$, (i) ν_{t+1} dominates ν_t ; (ii) the distribution of know-how eventually settles at a long run distribution ν^* ; and (iii) this long run distribution is never actually reached ($\nu_t \neq \nu^*$ at any t).

The intuition behind Proposition 1 is this: For part (i), the distribution of technology can never **become** worse, **since** a firm knowing technology θ would never **implement a technology** inferior to θ . For part (ii), since the best technology is indexed by $\theta = 1$, the **distribution** of technology can never advance beyond the one in which all **firms** know $\theta = 1$. **Thus**, since the distribution of technology can never fall back, it must either approach the distribution corresponding to all knowing $\theta = 1$, or one inferior to it. For part (iii), why must the distribution of technology always increase? Suppose, to the contrary, that its advance **halts** at some date t . In comparison to $t-1$, the main difference is that the distribution of technology in use is more advanced at t . Since this makes learning new techniques easier -- recall the **final** restriction on Ψ -- any **firm** that tried to **learn** at $t-1$ but failed, would wish to keep trying, in which case learning could not halt as **assumed**.⁸

Proposition 1 states that if the distribution of know-how ever begins to improve-- as can safely be assumed-- it will always improve. In this case the qualitative implications for industry

⁷ This result **summarizes** Propositions 2-S in Jovanovic and MacDonald (1991). Since a formal statement and proof require **significant** additional notation, the full details are omitted here.

⁸ This argument applies to a firm **having** some given information state in both periods. Since the learning technology is such that learning something new cannot be assured, there is always a positive fraction of firms that fail to learn, and so occupy the same information **state** during both periods.

output and price that stem from the evolution of the distribution of know-how and that are derived in the next 3 propositions. are predicted to hold at all stages of the industry's development.

The distribution of know-how improves over time, and under competition the implied cost reductions show up in a decline of the product price:

PROPOSITION 2: Ceteris **paribus**, the product price is lower at $t+1$ than at t ; i.e. for **any** x and ν , $P[x, \Phi(x, \nu)] \leq P(x, \nu)$.

PROOF: Using the **definition** of equilibrium, the **monotonicity** in θ of $q(\theta, x, \nu)$, and part (i) of Proposition 1, $P(x, \nu) = D[\int q(\theta, x, \nu) d\nu_t(\theta, x)] \geq D[\int q(\theta, x, \nu) d\nu_{t+1}(\theta, x)] = P[x, \Phi(x, \nu)]$. \parallel

Proposition 2 states that, given x , the price of output declines over time as supply shifts to the right and demand remains stationary. More generally, however, price can temporarily rise as a result of demand and supply shocks (X). The next result shows, however, that when the shocks are i.i.d., p_t declines in distribution. Let X_t be i.i.d., with distribution $F(\cdot)$. Denote the distribution function of p_t conditional only on ν , by $F^p(p | \nu) \equiv \int_{A(p, \nu)} dF(x)$, where $A(p, \nu) \equiv \{x | P(x, \nu) \leq p\}$. Then p_t is stochastically decreasing:

PROPOSITION 3: $F^p(\cdot | \nu_t)$ dominates $F^p(\cdot | \nu_{t+1})$.

PROOF: Since ν_t is stochastically increasing with t , and $q(\theta, x, \nu)$ is increasing in θ , $P(x, \nu_t)$ is declining in t for fixed x . It follows that $A(p, \nu_{t+1}) \supseteq A(p, \nu_t)$, and the claim follows. \parallel

Propositions 2 and 3 carry over to the evolution of industry output; for example, average output stochastically increases over time. But not every firm will experience continual output growth, even holding x constant: Firms whose technological growth falls far enough below the

industry average will produce less at $t+1$ than at t if the anticipated price decline occurs. Therefore, the cross-firm distribution of output may not always improve over time, since some **firms** will experience a fall in output. Nevertheless, a positive result can be obtained by “normalizing” the distribution of output: let normalized output be $\bar{q}(\theta, x, \nu) \equiv q(\theta, x, \nu)/q(\underline{\theta}, x, \nu)$, where $\underline{\theta}$ is the smallest value of θ in the support of ν_0 . Since $q(\theta, x, \nu)$ is increasing in θ , normalized output is distributed over $[1, \infty)$. Let ζ be the price elasticity of the firm’s supply curve.

PROPOSITION 4: If ζ is not increasing in θ , then for **fixed** x , the distribution of $\bar{q}(\theta, x, \nu)$ stochastically improves over time.

PROOF: For any q' , let $\bar{\theta}(p, x, q')$ be the solution for θ to $\bar{q}(\theta, x, p) = q'$. If ζ is not increasing in θ , then $\bar{\theta}$ is increasing in p . The fraction of **firms** at t having normalized output at most equal to q' is $\nu_t[\bar{\theta}(p_t, x, q')]$. But, since ν_{t+1} dominates ν_t , $\nu_t[\bar{\theta}(p_t, x, q')] \geq \nu_{t+1}[\bar{\theta}(p_t, x, q')]$. This exceeds $\nu_{t+1}[\bar{\theta}(p_{t+1}, x, q')]$ because of proposition 2, and because $\bar{\theta}$ is increasing in p . ||

The intuition **behind** Proposition 4 is this: As ν_t evolves to ν_{t+1} , two things happen. First, the output of **firms** that learned at t rises relative to that of the smallest **firms**, because some small firms will not have learned at t , and thus will not raise output. This output growth is tempered by the downward pressure on the product price needed to clear the product market. However, given the elasticity condition, the response of larger **firms** is proportionally no greater than that of smaller **firms**, so their normalized output in fact rises as price falls. With size measured by normalized output, this generates a declining proportion of small firms; that is, an improved distribution of normalized output.⁹

⁹ The elasticity condition is easily met. For example $c^q(q, \theta, x) = q^\alpha \bar{c}(\theta, x)$, for $\alpha > 1$, yields a price elasticity that is independent of θ . Further, the result’s simplicity stems from the absence of **fixed** costs, so that the smallest **firms** are always those who fail to learn anything at all.

Learning and Firm Growth: Propositions 2-4 were driven by the improvement of ν over time. In contrast, the next two propositions rely on how the firm's current know-how interacts with its learning opportunities. There is a basic tension: Technological laggards may have a greater incentive to improve through imitation, but the leaders may find it easier to learn still more through innovation. Thus whether higher θ leads to greater or lesser learning effort is unclear.

The next proposition deals with a case in which laggards will improve their technologies more often than leaders. The following condition implies that greater θ reduces learning effort: for $\theta' \geq \theta$, the learning technology is $\Psi(\theta' | l, \theta, x, \nu) \equiv [1 - \Psi^1(l, x, \nu)] + \Psi^1(l, x, \nu)\Psi^2(\theta' | x, \nu)$, where the range of Ψ^1 is $[0, 1]$ and for any x and ν , Ψ^2 is a distribution function; for $\theta' < \theta$, $\Psi(\theta' | \cdot) = 0$, as before. This restriction breaks learning into a two-step process. Step one determines whether the firm gets a technological draw, the probability of success being $\Psi^1(l, x, \nu)$, a function of learning effort but not know-how. In step two, if the firm does get a technological draw, it comes from $\Psi^2(\theta' | x, \nu)$, which depends neither on effort nor know-how. In this setup, sampling of new technology is as easy for laggards as it is for leaders. Of course, the firm will reject any draw less than θ , and since the leaders have larger B 's their returns from technological sampling are less.

Since greater know-how does not make learning easier or cheaper, laggards, who have more to gain from learning, will try harder and succeed more often:

PROPOSITION 5: The likelihood of learning a better technique, $\Psi^1[l(\theta, x, \nu), x, \nu][1 - \Psi^2(\theta | x, \nu)]$, is decreasing in θ .

PROOF: Since Ψ^2 is increasing in θ , it suffices to show that $\Psi^1[l(\theta, x, \nu), x, \nu]$ is decreasing in θ . To do so, let $\theta^0 < \theta^1$ and define $c^0 \equiv c^l[l(\theta^0, x, \nu), x, \nu]$, $\Psi^1(\theta^0) \equiv \Psi^1[l(\theta^0, x, \nu), x, \nu]$ and $\Psi^2(\theta^0) = \Psi^2(\theta^0 | x, \nu)$; define c^1 , $\Psi^1(\theta^1)$ and $\Psi^2(\theta^1)$ analogously. Writing ν' in place of $\Phi(x, \nu)$, that a firm knowing θ^0 selects $l(\theta^0, x, \nu)$ in preference to $l(\theta^1, x, \nu)$ implies

$$c^1 - c^0 \geq \beta [1 - \Psi^2(\theta^0)] [\Psi^1(\theta^1) - \Psi^1(\theta^0)] \left\{ \int_{\theta^0}^{\theta^1} \frac{V(\theta', x, \nu') d\Psi^2(\theta' | x, \nu)}{1 - \Psi^2(\theta^0)} - V(\theta^0, x, \nu') \right\}.$$

In conjunction with the analogous inequality for a firm knowing θ^1 , rearrangement gives

$$(*) \quad [\Psi^1(\theta^1) - \Psi^1(\theta^0)] \left\{ [1 - \Psi^2(\theta^1)] \left[\int_{\theta^0}^{\theta^1} \frac{V(\theta', x, \nu') d\Psi^2(\theta' | x, \nu)}{1 - \Psi^2(\theta^1)} - V(\theta^1, x, \nu') \right] \right. \\ \left. - [1 - \Psi^2(\theta^0)] \left[\int_{\theta^0}^{\theta^1} \frac{V(\theta', x, \nu') d\Psi^2(\theta' | x, \nu)}{1 - \Psi^2(\theta^0)} - V(\theta^0, x, \nu') \right] \right\} \geq 0.$$

The factor in braces in (*) is equal to

$$- \int_{\theta^0}^{\theta^1} V(\theta', x, \nu') d\Psi^2(\theta' | x, \nu') + V(\theta^0, x, \nu') - V(\theta^1, x, \nu') + \Psi^2(\theta^1) V(\theta^1, x, \nu') - \Psi^2(\theta^0) V(\theta^0, x, \nu').$$

Integrating the first term by parts and simplifying yields

$$\int_{\theta^0}^{\theta^1} \Psi^2(\theta' | x, \nu) dV(\theta', x, \nu') + V(\theta^0, x, \nu') - V(\theta^1, x, \nu').$$

Since $\Psi^2 \leq 1$, this expression is at most

$$\int_{\theta^0}^{\theta^1} dV(\theta', x, \nu') + V(\theta^0, x, \nu') - V(\theta^1, x, \nu'),$$

which is equal to zero. Thus the factor in braces in (*) is negative, in which case (*) yields $\Psi^1[l(\theta^0, x, \nu), x, \nu] \geq \Psi^1[l(\theta^1, x, \nu), x, \nu]$. ||

Proposition 5 implies that smaller firms will learn more frequently, but not, however, that they will necessarily grow faster; this is because the secular decline in price could affect them more than it affects big firms. Something more must be assumed to guarantee faster growth for smaller firms, and this is done in the next proposition. Given x , growth for a firm that knows θ is the random variable $q[\theta', x, \Phi(x, \nu)]/q(\theta, x, \nu) - 1$; denote its distribution by $G(g|\theta, x, \nu)$.

PROPOSITION 6: If ζ is not decreasing in θ , then for fixed x , $G(g|\theta, x, \nu)$ is stochastically decreasing in θ .

PROOF: Let $\theta^1 > \theta^0$ and $g = q[\theta^1, x, \Phi(x, \nu)]/q(\theta^1, x, \nu) - 1$; $g \leq 0$.

i) Let $g' > g$ and define $\underline{\theta}(g', \theta, x, \nu)$ to be the value of θ' solving $q[\theta', x, \Phi(x, \nu)] - g'q(\theta, x, \nu) = 0$; $\underline{\theta}$ is increasing in θ . Then $1 - G(g'|\theta^1, x, \nu) = \Psi^1[l(\theta^1, x, \nu), x, \nu][1 - \Psi^2[\underline{\theta}(g', \theta^1, x, \nu)|x, \nu]] \leq \Psi^1[l(\theta^0, x, \nu), x, \nu][1 - \Psi^2[\underline{\theta}(g', \theta^0, x, \nu)|x, \nu]] = 1 - G(g'|\theta^0, x, \nu)$. The inequality follows because $\Psi^1[l(\theta, x, \nu), x, \nu]$ is decreasing in θ (proof of Proposition 5) and $\underline{\theta}$ is increasing in θ .

ii) Let $g' = g$. Then $\underline{\theta}(g', \theta^1, x, \nu) = \theta^1$ and $G(g'|\theta^1, x, \nu)$ is the probability that a firm with know-how θ^1 fails to learn at t . Under the restriction on ζ , the probability that a firm with know-how θ^0 grows by as little as g' is either 0 (whenever $\Phi(x, \nu)$ differs from ν), or equal to the probability of failing to learn; the latter, according to Proposition 5, is rising in θ .¹⁰ Thus g' is the minimum in the support of $G(g|\theta^1, x, \nu)$ and below the minimum in the support of $G(g|\theta^0, x, \nu)$, and $G(g'|\theta^0, x, \nu) \leq G(g'|\theta^1, x, \nu)$. \parallel

The relation between Propositions 4 and 6 requires discussion. First, both results can

¹⁰ $q[\theta^1, x, \Phi(x, \nu)]/q(\theta^1, x, \nu) - q[\theta^0, x, \Phi(x, \nu)]/q(\theta^0, x, \nu) \leq 0$ is equivalent to $q[\theta^1, x, \Phi(x, \nu)]/q[\theta^0, x, \Phi(x, \nu)] - q(\theta^1, x, \nu)/q(\theta^0, x, \nu) \leq 0$, which follows immediately from the condition on ζ and $P[x, \Phi(x, \nu)] \leq P(x, \nu)$.

hold **simultaneously** since both admit the case in which the elasticity of supply is **independent** of θ . In this case the distribution of normalized output becomes less concentrated **over** time while small firms have a greater tendency to grow. This occurs because the tendency for small **firms** to grow does not eliminate the fact that some small firms fail to learn, and fall behind,

Propositions 5 and 6 are driven by the assumption that all **firms** sample from the same technological pool. But some cost-reducing improvements are incremental -- they build directly on technology in place-- as opposed to more fundamental improvements based on **discoveries** that are either new to all or new to the industry. When improvements are incremental, technological leaders sample new technology from a **better** pool than laggards do. The assumption that **increments** are sampled seems especially appropriate when it comes to technological improvements introduced by current leaders. But it also makes sense for advances achieved through imitation: A **firm** that is sufficiently far behind may well **find it harder** to imitate an advanced technology than will a **firm** whose know-how is closer to the technology it is trying to imitate. If learning is indeed of an “incremental” nature in this sense, and if this effect is strong enough, then in contrast to the assertions of propositions 5 and 6, higher θ can raise the incentive to acquire new know-how. This effect is present in the example at the end of the paper, in which there are but three technologies: low, medium and high. The invention process endows medium tech firms with an advantage -- through innovation they can become high tech much easier than low tech **firms** can. As a result, in comparison to low tech **firms**, medium tech **firms** devote more effort to innovation during the entire life-span of the industry; Even though they have less to learn, and imitate less vigorously, they are more likely to succeed in implementing better technology.

Propositions 5 and 6 continue to hold even if neither Ψ^1 nor Ψ^2 depends on ν . That is, they hold **even if imitation is impossible**.. These two propositions are driven entirely by the cross-sectional fishing out effect. Thus there are two distinct forces in the model that push the population of **firms** towards technological convergence: (a) the diffusion of technology that results from the laggards’ efforts to imitate the leaders, and (b) the fishing out effect that causes

the laggards to search harder. The first force likely shows up in any industry, while the second arises only whenever technological laggards **are** about as efficient in finding better technologies as the leaders are, as is more **likely** when new technology is not closely related to old technology.

Optimality. Unless Ψ does not depend on ν , the distribution of know-how in use affects how **firms** learn- an externality; thus a social optimum is not likely to coincide with equilibrium. The next result shows that there is a sense in which a social planner would prefer that firms engage in more of **all** information gathering activities. *Some* of these activities may be **imitative** in the sense that distribution of returns to learning efforts depends on others' know-how, summarized by ν . **Other** information gathering activities may have a return that does not depend on ν , and in this sense it is **innovative**.

Suppose that a social planner ranks outcomes by consumers' plus producers' surplus. **Equilibrium** surplus is

$$E \left\{ \sum_{t=0}^{\infty} \beta^t \left[\int_0^{Q(x_t, \nu_t)} D(z, x_t) dz - \int c[q(\theta, x_t, \nu_t), l(\theta, x_t, \nu_t), \theta, x_t] d\nu_t(\theta) \right] \right\},$$

where $Q(x_t, \nu_t) \equiv \int q(\theta, x_t, \nu_t) d\nu_t(\theta)$ is industry output at t . Consider a **nonzero** vector of learning efforts l' . Suppose that at a date t' , **firms'** learning efforts are $l(\theta, x_t, \nu_t) + \epsilon l'$, for $\epsilon > 0$, but that **firms** follow the equilibrium policy otherwise. Let $W(\epsilon)$ denote the surplus generated by this behavior.

PROPOSITION 7: Unless Ψ does not depend on ν , $dW(\epsilon)/d\epsilon > 0$ at $\epsilon = 0$. That is, the planner prefers more learning efforts at t' .¹¹

¹¹ The proof of Proposition 7 is in Jovanovic and MacDonald (1991). The argument **requires** conditions guaranteeing differentiability of W . A related result is in Rustichini and Schmitz (1991).

In practice, the main policy tool employed to encourage innovative activity is patent protection. Patents encourage innovation by stifling imitation. From **this perspective** Proposition 7 may seem counter-intuitive in that it implies not just that innovation is too low in equilibrium, but that imitation is too **low** as well; i.e. deterring imitation is, in fact, detrimental. This occurs because the result does not compare equilibrium to an alternative in which the planner must intervene solely via a patent system. Rather, the planner has a richer set of policy tools that **influence** both imitation and innovation, encouraging both more discoveries and faster spread of what is found. Of course, equilibrium and optimum coincide if learning is independent of others' know-how, i.e., if Ψ does not depend on ν .

This result suggests caution in the design of policies whose goal is to encourage cost reduction efforts: If it is possible **simultaneously** to provide incentives for innovation and imitation, then imitation is not necessarily a bad thing. The result is limited in two respects, however. First, it does not compare equilibrium with the planner's **global** optimum; this comparison will be made in an example, below. Second, the result states that the planner would prefer a small increase in all learning activities, but it does not say what the planner's favorite small change would look like.

The Lone Run. Is the best technology ($\theta = 1$) eventually uncovered? Will every firm **eventually** learn $\theta = 1$? Or will firms be different forever?" The answers depend on the learning technology and **cost function**. If the marginal cost of learning is always positive, then, at most a fraction of firms will ever use the best technology -- not **all** firms will get it. After a point, firms have learned enough that the remaining scope for cost reduction becomes too small to justify the effort needed to replace an existing method with the best one.

The above scenario admits the possibility that all firms will converge to a $\theta < 1$ and stay there forever. But if, in addition, the learning technology is such that given θ the firm

¹² Detailed answers to these question are provided in Jovanovic and MacDonald (1991).

might, in one step, learn any **greater** θ (i.e. for all θ , the support of $\Psi(\theta' | l, \theta, x, \nu)$ is $[\theta, 1]$), then heterogeneity persists forever. That is, when the **learning** technology can yield a diverse set of new techniques, not only do some firms never use the **best** technology, but also there is no technology ultimately used by all **firms** -- know-how must remain dispersed. Since the output of a **firm** and its value are both positively related to its know-how, long run dispersion of **know-how** implies long run dispersion of output and **firm** values. This means that even for mature industries defined by narrowly defined commodities, the distribution of output should not be concentrated, even though **firms** can imitate one another.

II. Innovation versus **Imitation**.

This section explores innovation and imitation in two examples of the general model that illustrate the properties of equilibrium and of the social optimum. The second example is then used to interpret the **diffusion** of diesel locomotives in the U.S. Neither example will have aggregate risk and so X_t does not appear below.

The general model allowed for a vector of learning activities, l . In the examples there will be just two ways to learn -- innovation and imitation. Innovative effort η gives the **firm** a draw θ' from the distribution $N(\theta' | \theta)$ with probability η . Observe that N depends on the firm's own know-how, but *not* the know-how of others. This is the sense in which effort η is innovative. Likewise, imitative effort, μ , gives the **firm** a draw from a distribution $M(\theta' | \nu)$ that dominates ν and improves whenever ν does; $M(\cdot | \nu) = \nu$ is an example. Effort μ is imitative since ν represents what others know and the draw does not depend on the firm's own know-how. M will dominate ν **if, for** example, the **firm** can direct its imitative effort towards the leaders. Assuming innovative and imitative luck are independent, $\Psi(\theta' | \eta, \mu, \theta, \nu) = [1 - \eta + \eta N(\theta' | \theta)][1 - \mu + \mu M(\theta' | \nu)]$ for $\theta' \geq \theta$, and $\Psi(\theta' | \eta, \mu, \theta, \nu) = 0$, otherwise.

Example 1: Three Technologies. Assume that *i*) there are three technologies -- $\theta_0 = 1$, $\theta_2 = 5$ and $\theta_3 = 15$ -- with all firms knowing only θ_0 at $t=0$; *ii*) costs are $c(q, \eta, \mu, \theta) =$

$\frac{1}{2}q^2/\theta + \frac{1}{2}\eta^2 + \frac{1}{3}\mu^2$; iii) imitation is undirected: $M(\cdot | \nu) = \nu$; iv) innovation is such that given a draw from N , a firm knowing θ_0 learns θ_1 with probability .05 and θ_2 with probability .01, and a firm knowing θ_1 learns θ_2 with probability .05; v) the discount factor is $\beta = .98$; vi) demand is $D(Q) = 2 - 2.5Q$. Let ν_i denote the fraction of firms knowing θ_i at t . Then the probability with which a firm knowing only θ_0 learns θ_1 is $.05\eta + \mu\nu_{1t} - .05\eta\mu\nu_{1t}$; that is, the probability of either innovation or imitation yielding θ_1 , minus the probability of both doing so. Expressions for the probabilities of other transitions are analogous.

Figure 2 charts the industry's evolution. In Panel a) all firms start out with low tech know-how, θ_0 . Innovation soon yields the discovery of medium tech know-how, θ_1 , which diffuses quickly due to imitation. High tech know-how, θ_2 , is also discovered early-- since there are many firms, any of which might uncover high tech-- but its diffusion lags behind the diffusion of medium tech; this occurs because high tech spreads more easily, via imitation, once medium tech achieves wide use. Eventually high tech swamps the less efficient techniques. Panels b) and c) display innovation and imitation effort for a low or medium tech firm. Initially, since there are few firms to imitate, imitation effort is nil and all advance is due to innovation. But as innovation breeds heterogeneity in know-how, the return to imitation rises rapidly and imitation soon substitutes for innovation. Observe that, in comparison to medium tech, low tech firms devote greater effort to imitation and less to innovation. This occurs because medium tech firms have only the high tech population to imitate, while low tech firms can learn from any medium or high tech firm. In addition, the cost saving low tech firms realize by learning is greater than the saving realized by medium tech, which explains why the difference in imitation never disappears entirely.

Innovation and imitation are substitutes, and this can complicate empirical work on the effects of R & D. For example, suppose that R & D data are primarily measures of innovative efforts. Then regressions of industry output growth on R & D expenditures will typically *understate* the influence of R & D on output growth; this occurs because the substitution relationship between innovation and imitation tends to cause them to be *negatively correlated*,

in which case the familiar omitted variable bias argument leads to the conclusion that the estimated effect of R & D on growth will be biased downward. In the example, the correlation of aggregate innovation expenditures with imitation expenditures is **-.23**, and a regression of the growth rate in output on innovation expenditures and a constant yields a slope coefficient of **.53**; including imitation expenditures in the regression raises this **figure** to **.61**.

Since **firms** can learn from one another, and learning leads to output growth, growth is fastest when the scope for learning -- differences in know-how -- is greatest. Thus it should be *dispersion* in output, not its mean, **that** drives imitation and raises growth. Figures **2(d)** and **3(d)** show that equilibrium and optimum both involve a positive correlation between the growth of the industry's output, and the variance of output among **firms**. The *growth* rate of output falls as the industry evolves, in agreement with observation by **Gort** and **Klepper** (1982). This generates a negative correlation **between** growth and level of output; however, growth and variance in output are positively correlated since imitation occurs in response to variance in know-how and results in greater subsequent output. Figures 2(e) and 3(e) display the declining price paths implied by equilibrium and optimum.

Figure 3 depicts diffusion, innovation and imitation in a social optimum. In this example the gains from improved know-how are large, and the dramatic difference between equilibrium and optimum reflects this. New know-how is discovered *more* quickly and spreads faster, as Proposition 7 would suggest. Indeed, the gains to getting high tech are so great that medium tech never gains widespread use. Instead, great effort is spent on imitation and high tech spreads quickly.

Example 2: Two Technologies: Assume that **i)** there are only two values of θ -- low tech, θ_0 and high tech, θ_1 . **ii)** costs are again quadratic, but more general -- $c(q, \eta, \mu, \theta) = c_q q^2 / 2\theta + \frac{1}{2} c_\eta \eta^2 + \frac{1}{2} c_\mu \mu^2$; **iii)** instead of being downward sloping, demand is perfectly elastic at price p_0 ; **iv)** given innovation effort η , the probability with which a low tech firm gets high tech know-how via innovation is $\eta\delta$, where $\delta \in (0, 1)$; and **v)** given imitation effort μ and

the fraction ν_t of firms knowing high tech at t , the probability with which imitation yields high tech know-how is $\mu\sqrt{\nu_t}$.¹³ This example resembles models for the transmission of disease in which ν would be the fraction of the population that is sick. In contrast to those models, contagion is endogenous.

Let the *net* benefit to learning be $\Delta(\nu) \equiv V[\theta_1, \Phi(\nu)] - V[\theta_0, \Phi(\nu)]$; Δ is falling in ν since imitation becomes easier as the number of high tech firms rises, raising $V[\theta_0, \Phi(\nu)]$, and $V[\theta_1, \Phi(\nu)]$ is independent of ν because demand is infinitely elastic. The *first* order conditions characterizing a low tech firm's choice of η and μ are

$$-c_\eta \eta + \beta \delta (1 - \mu \sqrt{\nu_t}) \Delta(\nu) = 0$$

$$-c_\mu \mu + \beta \sqrt{\nu_t} (1 - \eta \delta) \Delta(\nu) = 0.$$

It is not hard to check that if $\sqrt{\nu} \cdot \Delta(\nu)$ is increasing in ν , the firm's innovation effort declines as ν grows, while its imitation effort rises. In what follows, this property will be assumed.¹⁴ Intuitively, $\Delta(\nu)$ is the net benefit associated with obtaining high tech know how, which declines over time; this encourages less of both methods of learning, a scale effect of a sort. However, that imitation is becoming easier promotes a substitution of imitation for innovation. Thus, the assumption made here is that the substitution effect dominates the scale effect.

The variety of industry evolutions is illustrated by two polar cases. In "pure imitation",

¹³ The fraction of high-tech firms is a sufficient statistic for the distribution of technology over firms, and that is why the symbol ν is borrowed to denote this fraction here. This specification of imitation assumes imitation by low tech firms can be partially directed towards high tech firms; i.e. given that imitation effort has turned up some firm, the probability that the firm is high tech is $\sqrt{\nu_t} > \nu_t$ (provided $\nu_t \neq 0$).

¹⁴ This property must hold over time in the sense that $\nu_0 \Delta(\nu_0) = 0$, while ν_t is rising over time and $\lim_{t \rightarrow \infty} \sqrt{\nu_t} \cdot \Delta(\nu_t) > 0$.

innovation is almost impossible -- $\delta = 0$ -- and evolution is driven by **imitation**.¹⁵ In “pure innovation”, imitation is ruled out. These cases represent the extremes in terms of the importance of informational linkages; they have distinct implications for, among other things, the diffusion of technology, the distribution of output, and the productivity of R&D.

A. Diffusion of new technology. From the first order conditions, the rate of adoption of high tech know-how is $v_{t+1} - v_t = \beta(1-v_t)v_t\Delta(v_t)$, which is small both early in the industry’s evolution, when $v \approx 0$, as well as much later, when $v \approx 1$; otherwise, diffusion occurs more quickly, and in this sense, pure imitation must result in the familiar “S-shaped” diffusion pattern. In fact, since A is declining in v , the maximal rate of diffusion occurs before a majority of firms use high tech know-how. In contrast, pure innovation implies that the net return to acquiring high tech know-how does not fall over time; this is because the composition of existing know-how does not influence the scope for learning.¹⁶ Thus innovative effort of low tech firms is constant over time, and $v_{t+1} - v_t = (\eta\delta)^{t+1}$, which is declining and concave in t . Thus under pure innovation, diffusion is quickest at the **outset**, while under pure imitation, diffusion is initially slow, then more rapid, and finally slow once again.

B. The time path of the distribution of output. Since price is constant over time, the outputs of high and low tech firms are constant over time and average output is simply a **rescaling** of v_t ; thus the results on diffusion apply immediately to mean output. The variance of output at t is **proportional** to $v_t(1-v_t)$. Thus given the diffusion paths just discussed, pure imitation yields

¹⁵ To allow high tech know how to emerge at all, $\delta > 0$ must be assumed. An alternative is to set $\delta = 0$ and endow a few firms with high tech know how right at the start. Also, the first version of this paper (Jovanovic and MacDonald, 1988) contains a variety of comparative dynamics results for the pure models.

¹⁶ The infinite elasticity of demand plays a role here too, but when comparing the two pure cases, it is the influence of the distribution of know how on learning possibilities that is central.

heterogeneity in output slowly, with the rapid imitation phase driving it out quickly; pure innovation results in a more rapid rise in heterogeneity and a more gradual decay.

C. The productivity of R & D spending. Under pure innovation there are no external effects in the learning technology. Only low tech firms try to progress, and their effort is constant. Therefore, the observed productivity of R&D spending (i.e. diffusion per dollar of expenditure) is a constant. Under pure imitation, on the other hand, the rise in the number of high tech firms makes it easier for the low-tech firms to copy them, causing productivity to rise over time.

Diffusion of the Diesel Locomotive

The twentieth century has seen a host of innovations in the Railroad industry, but all are dwarfed by the replacement of steam engines by diesels. This section interprets the data on the diffusion of diesels using the pure imitation model.”

The first usable diesel locomotive was invented by Rudolf Diesel in 1912. Diesels were first used in the U.S. in 1925, and by 1968 they had displaced steam engines **entirely**.¹⁸ Panel a) of Figure 4 displays diesels in use in the U.S. (1925-67) as a fraction of the total number of locomotives; the data are from the U.S. Bureau of the Census, *Historical Statistics of the United States*, series Q296-9. Since this fraction is steadily rising, the distribution of technology increases over time.

Spillovers cause the likelihood of switching technologies to depend on the distribution of technology in use. Indeed, in the pure imitation model, the “hazard” rate $h_t \equiv (\nu_{t+1} - \nu_t) / (1 - \nu_t)$ can be “backed out” of the diffusion *data*, and is predicted to be rising. Panel (b) of Figure 4 displays the hazard implied by the diffusion in panel (a). The hazard increases through most of

¹⁷ In Jovanovic and MacDonald (1988) the pure innovation model is used to study data on the diffusion of mechanized loading techniques in the U.S. underground coal industry.

¹⁸A few electric and “other” locomotives are ignored in what follows since, altogether, they never amounted to more than 2% of the total number of locomotives in use.

its range, and, in fact, **only** fails to do so when more than 99% of locomotives were diesels; in this case the denominator of the hazard-- I-v,-- is less than .01 and some erratic behavior for measured h_t , is not surprising. The pure imitation model has the stronger implication that imitation effort $\mu_t = h_t \sqrt{v_t}$ rises over time. Panel (c) displays μ_t , calculated from the diffusion data in panel (a). Imitation effort generally rises and only fails to do so when almost all locomotives in use are diesels.

While these illustrative calculations do not prove that an informational model underlies these data, it is worth noting that the main alternative hypothesis -- a vintage capital model-- leaves much unexplained: **First, new** steam locomotives were produced long after the introduction of diesels; see Interstate Commerce Commission (1950. Table A-4 and A-5). And second, there is no evidence of a bell-shaped distribution of ages of locomotives at the time of introduction of diesels; a bell-shape is key for a vintage explanation for the S-shape in Panel a). Nor did the substitution of diesels for locomotives merely reflect the cheapening of oil relative to coal. In fact, over the 1940-60 period, during which the primary displacement of steam engines occurred, the relative price of oil to coal rose by about 15%; see U.S. Bureau of the Census. *Historical Statistics of the United States*, Series M96, M139.

III. Conclusion.

This paper has analyzed competition among firms that differ only in terms of their productive knowledge. Diffusion of technology takes time as a result of informational barriers defining firms. Restrictions on observables were driven by this gradual spread of know-how.

Two key assumptions should be emphasized: (i) The informational unit is firm, and (ii) Informational barriers take time and effort to overcome.

The firm is a legal entity: Patents are granted to firms, and information-sharing mechanisms such as patent-swapping arrangements and licensing agreements are made among firms. On these grounds it makes sense to think of the owner of a piece of information -- information that other firms can try to acquire -- as a firm. Patents represent a barrier to the

flow of implemented information among firms. Moreover, when a firm's employees work on the same premises, they can and do share information among themselves differently and more often than they do with others. These factors point to the firm as the appropriate unit of analysis.

Important informational barriers may, however, exist within the firm, especially within large firms. **Holmstrom** (1982) analyzes incentive problems that arise within the **firm** -- problems that may deter a plant manager from sharing his technological know-how with his peers. Consistent with this idea, **Mansfield** (1963) has shown that the spread of a new technology within the firm can take almost as long as its spread within the industry. To explain diffusion lags within the firm this **model** must interpret them as resulting from informational barriers between decision units making up the **firm**. Now if plants or individuals are treated as the decision units, the model's predictions are about plants or individuals, not firms. Whatever the decision unit is, however, the results apply to the discovery and spread of know-how in a group of such decision units.

Although informational barriers among firms (and perhaps within firms too) seem to matter, how much is not yet clear. In particular, is the spread of technological know-how slow enough to explain much of the variance in **firm size** within industries, and in the observed timing of the adoption of new technologies? The answers hinge on how **easily firms** can imitate one another; indeed, if imitation were as easy as obtaining a blueprint or recipe, a theory focusing on institutional features like patents and licenses would be more relevant than the theory presented in **this** paper. But imitation is typically not **that** easy: using another's idea involves more than simply obtaining a blueprint. Just as mastering a subject demands more than buying a textbook. This explains why **firms** in fact classify most of their R & D expenditures as "applied", and why information lags probably are important for understanding how industries evolve.

References

- Andolfatto, David, and Glenn M. MacDonald, "Endogenous Technological Change, Growth, and Aggregate Fluctuations." revised, June 1993.
- Arrow, Kenneth, "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29 (1962). 155-173.
- Chari, V.V. and Hugo Hopenhayn, "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy* 99, no. 6 (December 1991): 1142-65.
- Flaherty, M. T., "Industry **Structure** and Cost-Reducing Investment," *Econometrica*, 48 (July 1980). 1187-1209.
- Gort, Michael, **and** Steven Klepper, "Time-Paths in the Diffusion of Product Innovations," *Economic Journal*, 92 (September 1982). 630-53.
- Holmstrom, **Bengt**, "Moral Hazard in Teams," *Bell Journal of Economics*. 12 (Autumn 1982), 324-340.
- Interstate Commerce Commission, Bureau of Transport Economics and Statistics, *Study of Railroad Motive Power*, Washington, D.C., 1950.
- Jovanovic. **Boyan**, and Glenn M. MacDonald, "Competitive Diffusion," 1988,
- Jovanovic. **Boyan**, and Glenn M. MacDonald, "Competitive Diffusion," Bradley Policy Research Center, Financial Research and Policy Studies Working Paper FR 92-08, December 1991.
- Jovanovic, **Boyan**, and **Rafael** Rob, "The Growth and Diffusion of Knowledge," *Review of Economic Studies* 56, no. 4 (October 1989): 569-82.
- Lucas, Robert, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22 (1988). 3-42.
- Mansfield. Edwin, "**Intra-firm** Rates of Diffusion of an Innovation, *Review of Economics and Statistics*, 65 (November 1963). 348-49.
- Mansfield, Edwin, Mark Schwartz and Samuel Wagner, "Imitation Costs **and** Patents: An Empirical Study," *Economic Journal* 91 (December 1981): 907-18.

Nabseth, Lars, and George Ray, *The Diffusion of New Industrial Processes: An International Study*, Cambridge, Cambridge University Press, 1971.

Pakes, Ariel and Mark Shankerman, "The Rate Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources," in Z. Griliches (ed.), *Patents, R&D, and Productivity*, University of Chicago Press, 1984.

Rogers, Everett M., *Diffusion of Innovations*. 3rd Ed., New York: Free Press, 1983.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy*, 98 (October 1990, supplement). S71-102.

Rustichini, Aldo and J. Schmitz, "Research and Imitation in Long Run Growth," *Journal of Monetary Economics*, 27 (April 1991), 271-292.

Schumpeter, Joseph A., *The Theory of Economic Development*, Cambridge: Harvard University Press, 1934.

Shell, Karl, "A Model of Inventive Activity and Capital Accumulation, in *Essays in the Theory of Optimal Economic Growth*, Karl Shell (ed.), Cambridge MIT Press, 1967.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39 (1957), 312-20.

Spence, A. M., "Cost Reduction, Competition, and Industry Performance," *Econometrica*, 52 (January 1984). 101-21.

Figure 1. Dynamic Random Access Memory Industry

a) Diffusion of Chip Density, 1973-1986

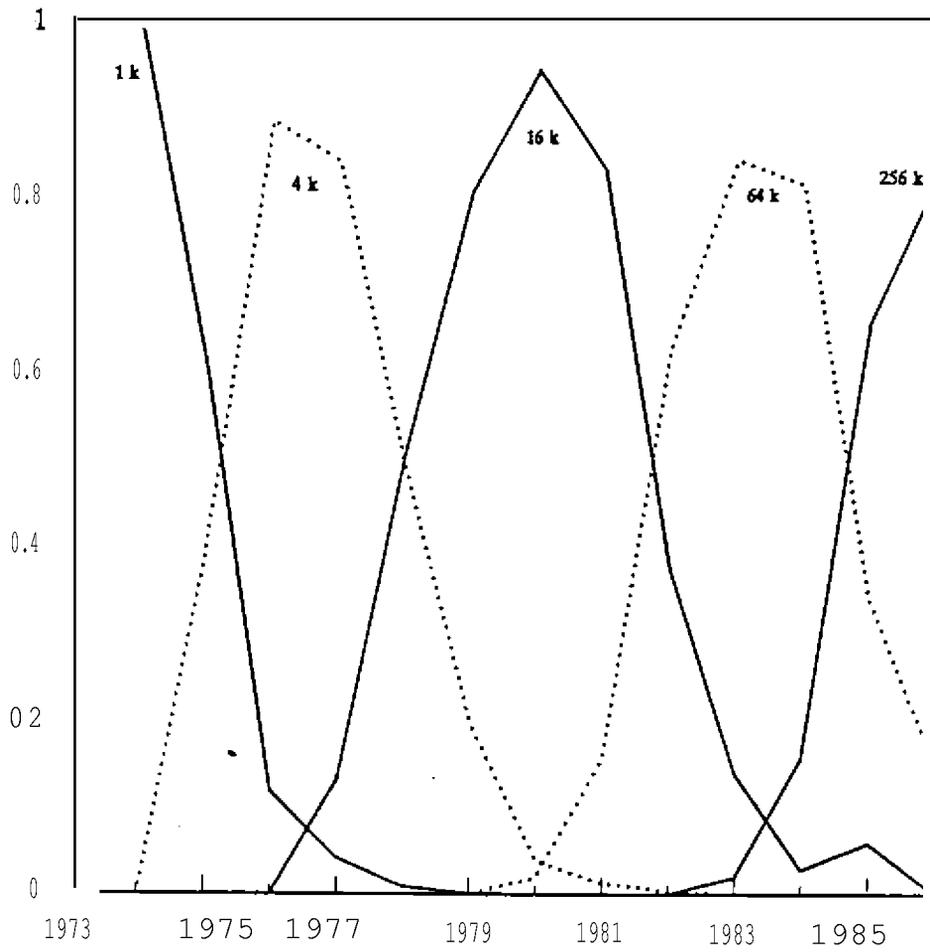


Figure 1. Dynamic Random Access Memory Industry

b) Industry Output, 1973.86 (Trillions of Bits)

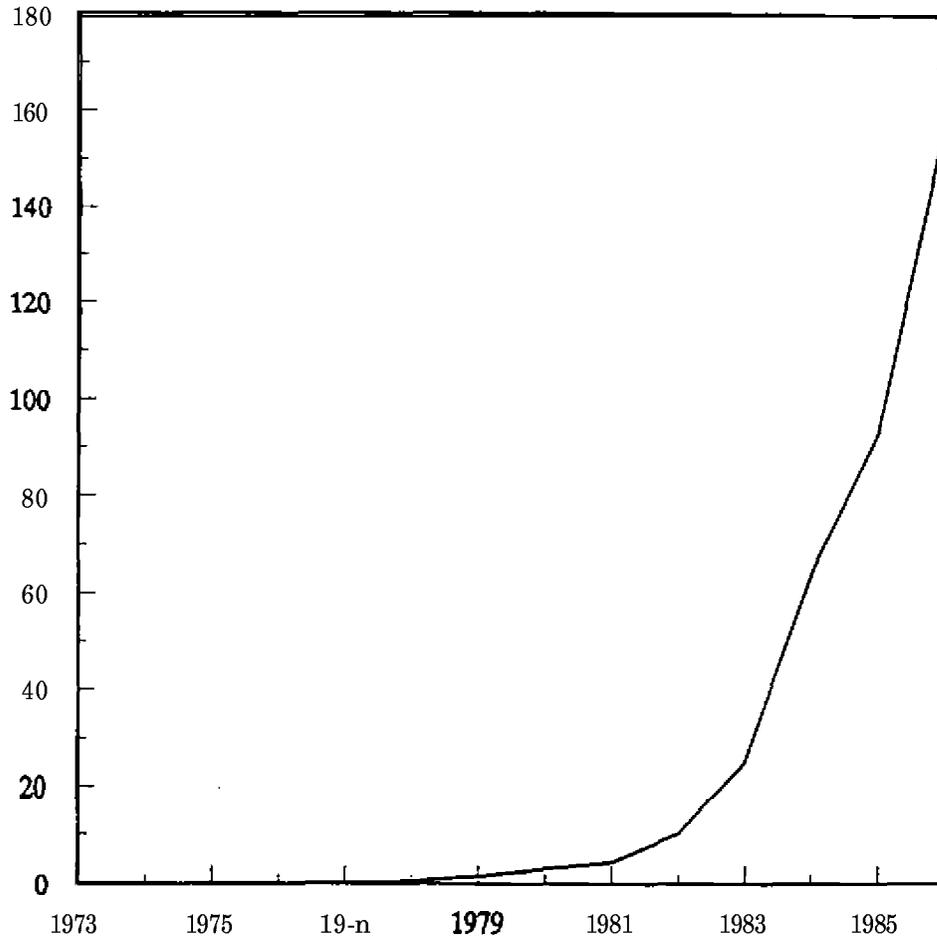


Figure 1. Dynamic Random Access Memory Industry

c) Price Per Bit, 1978-86 (Real Millicents per Bit)

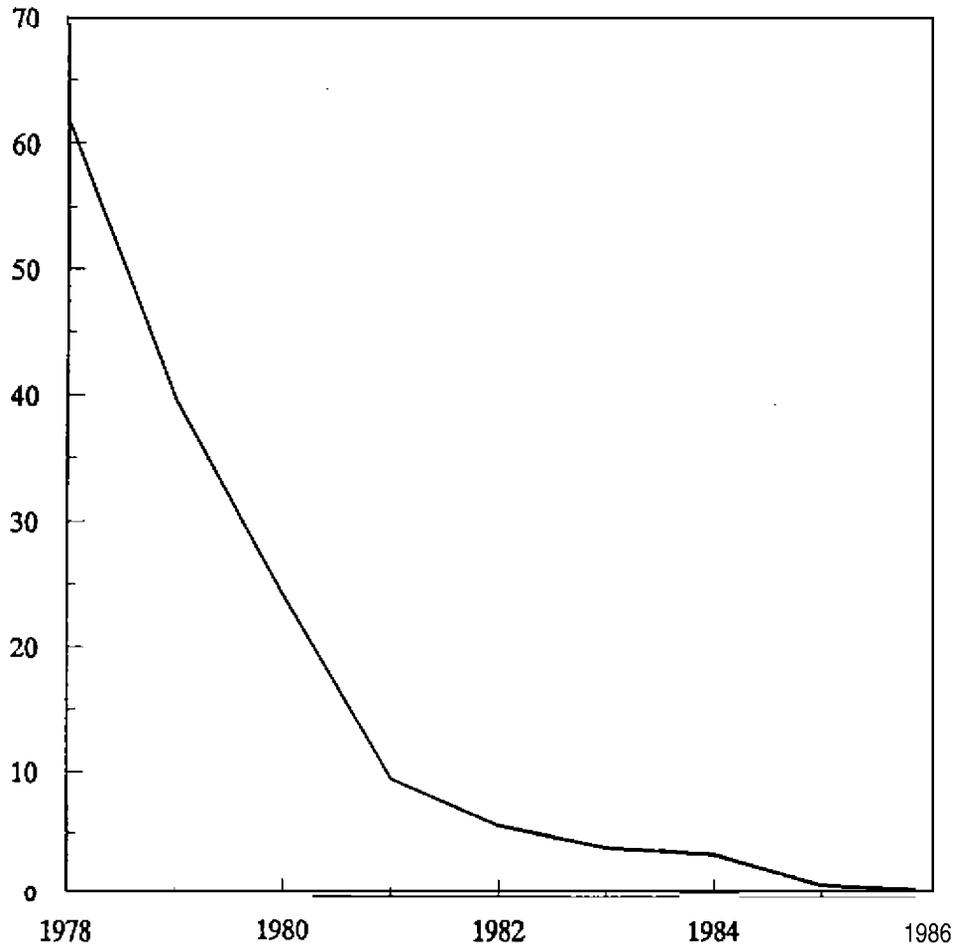


Figure 2. Example of Equilibrium

a) Diffusion

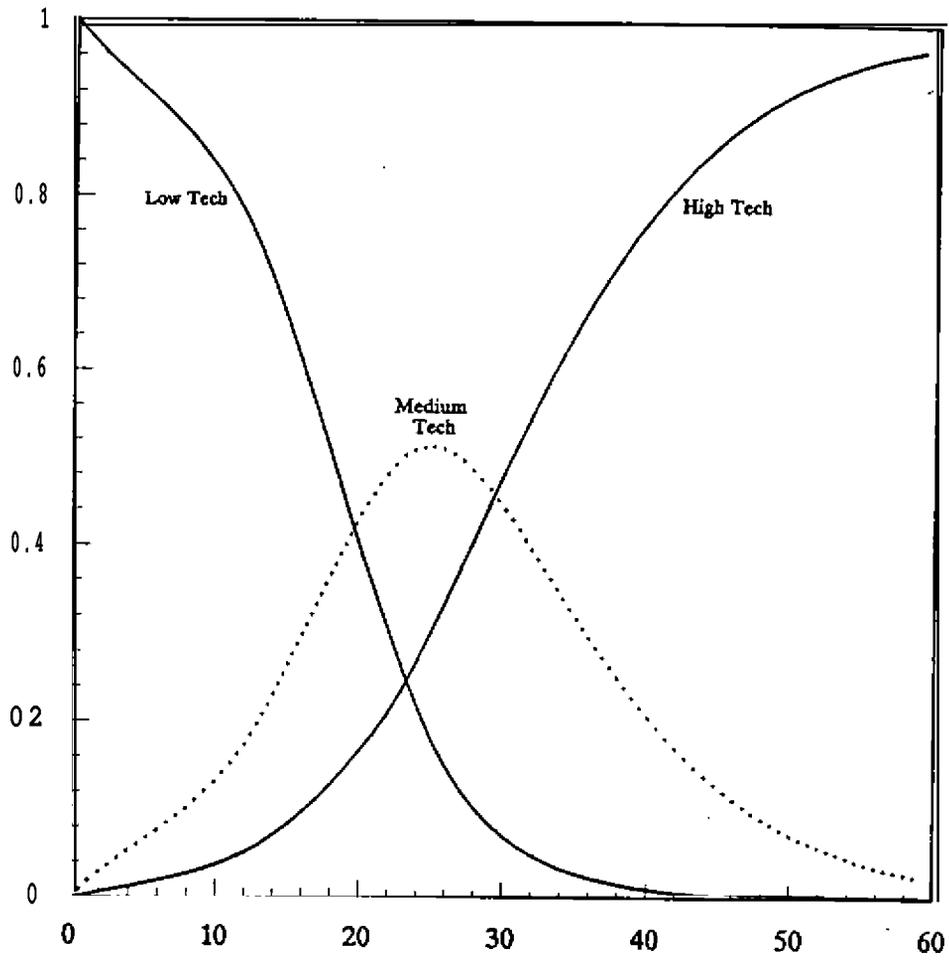


Figure 2. Example of Equilibrium

b) Innovation Effort (η)

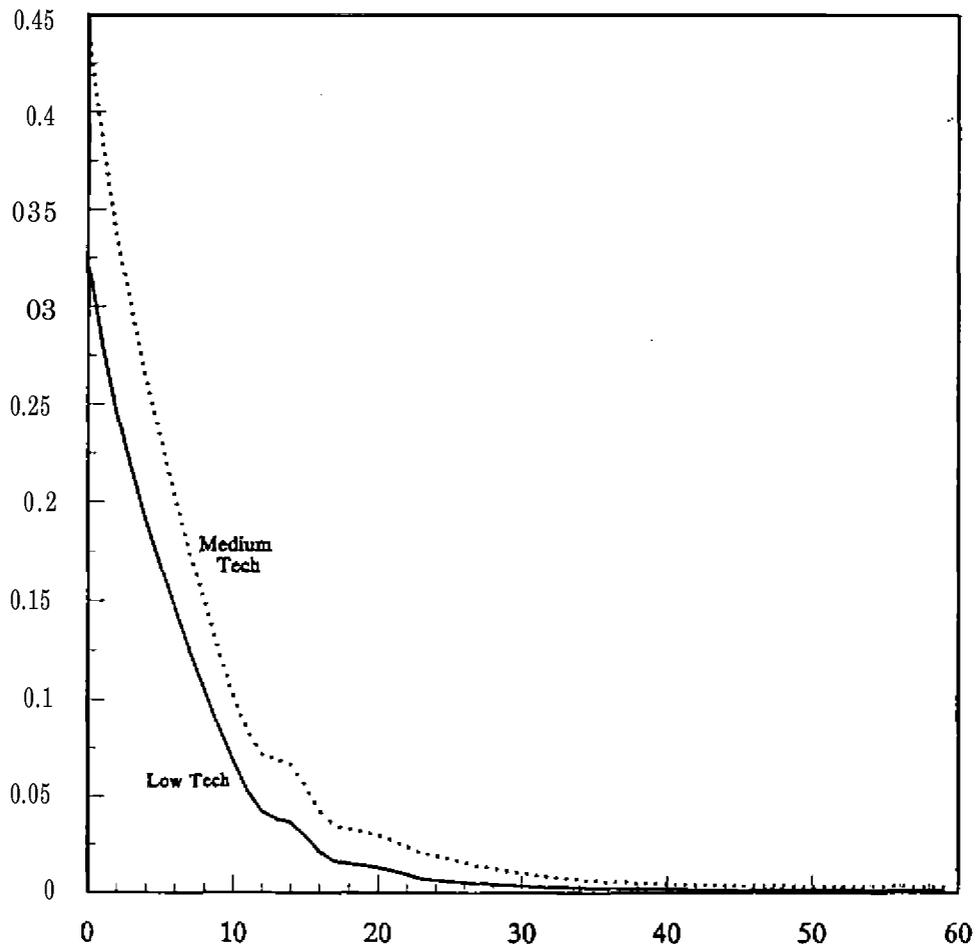


Figure 2. Example of Equilibrium

c) Imitation Effort (μ)

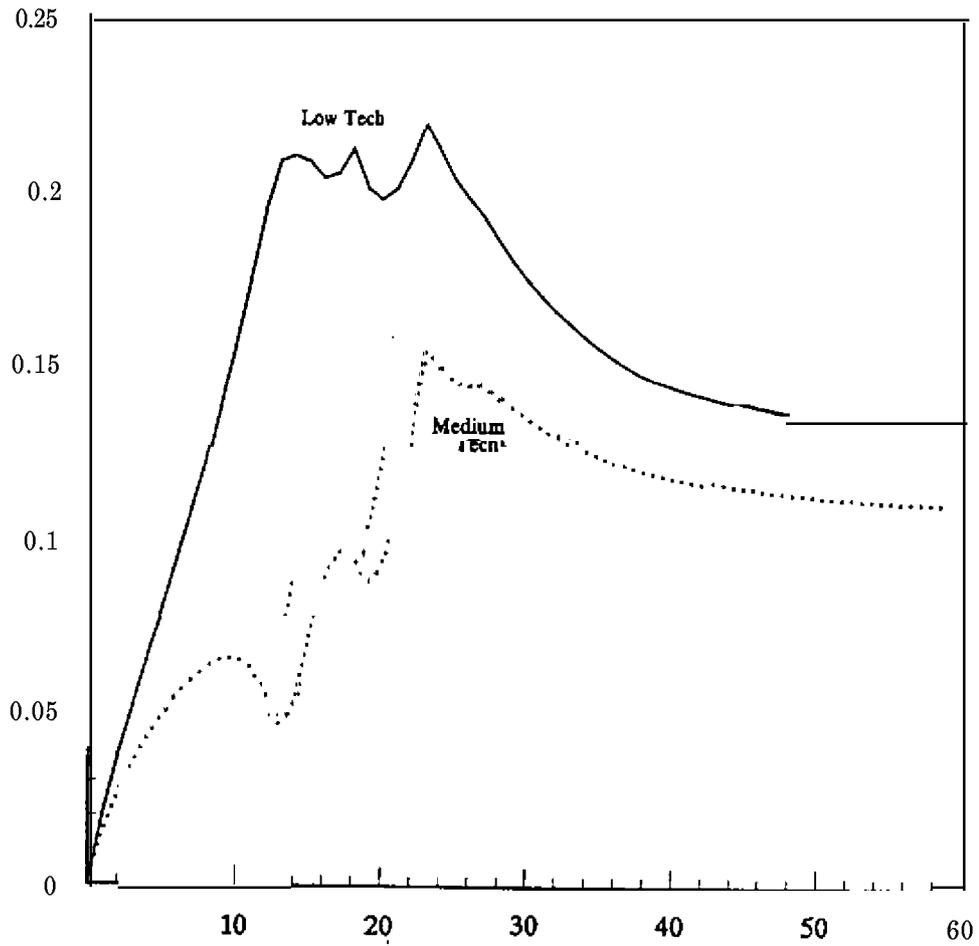


Figure 2. Example of Equilibrium

d) **Industry Output (Q)** and its Variance Among Firms

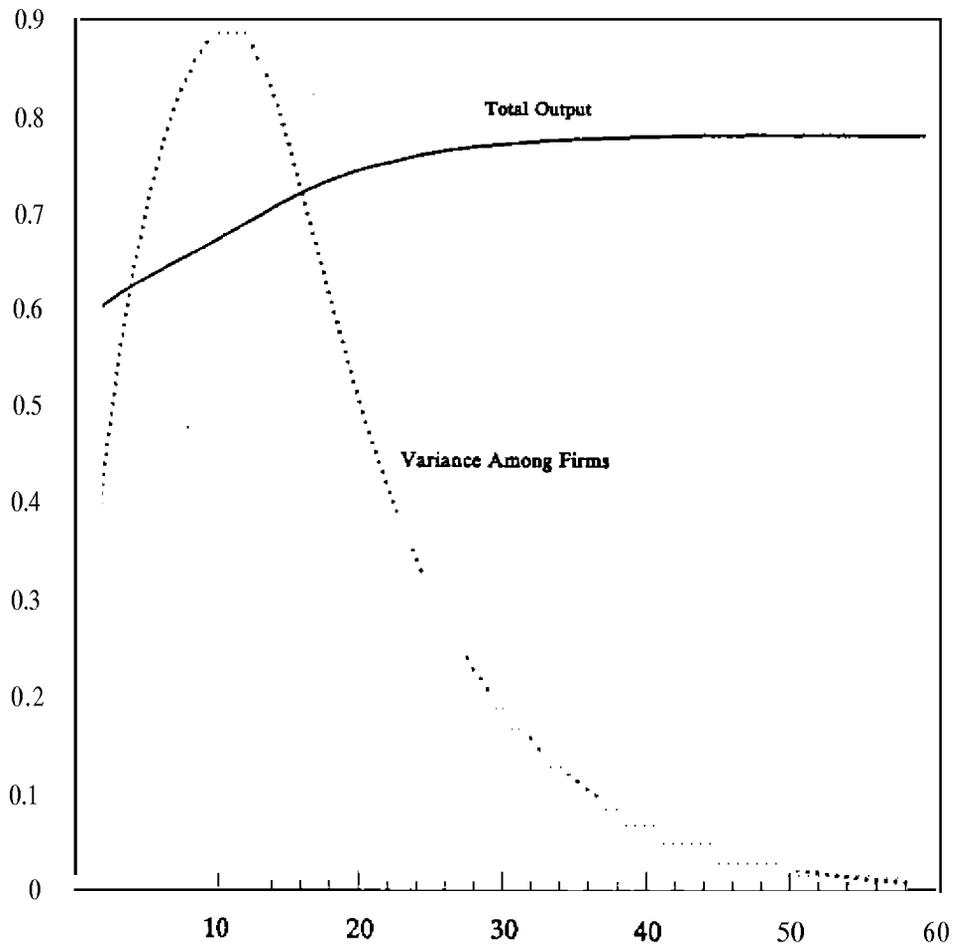


Figure 2. Example of Equilibrium

e) Price

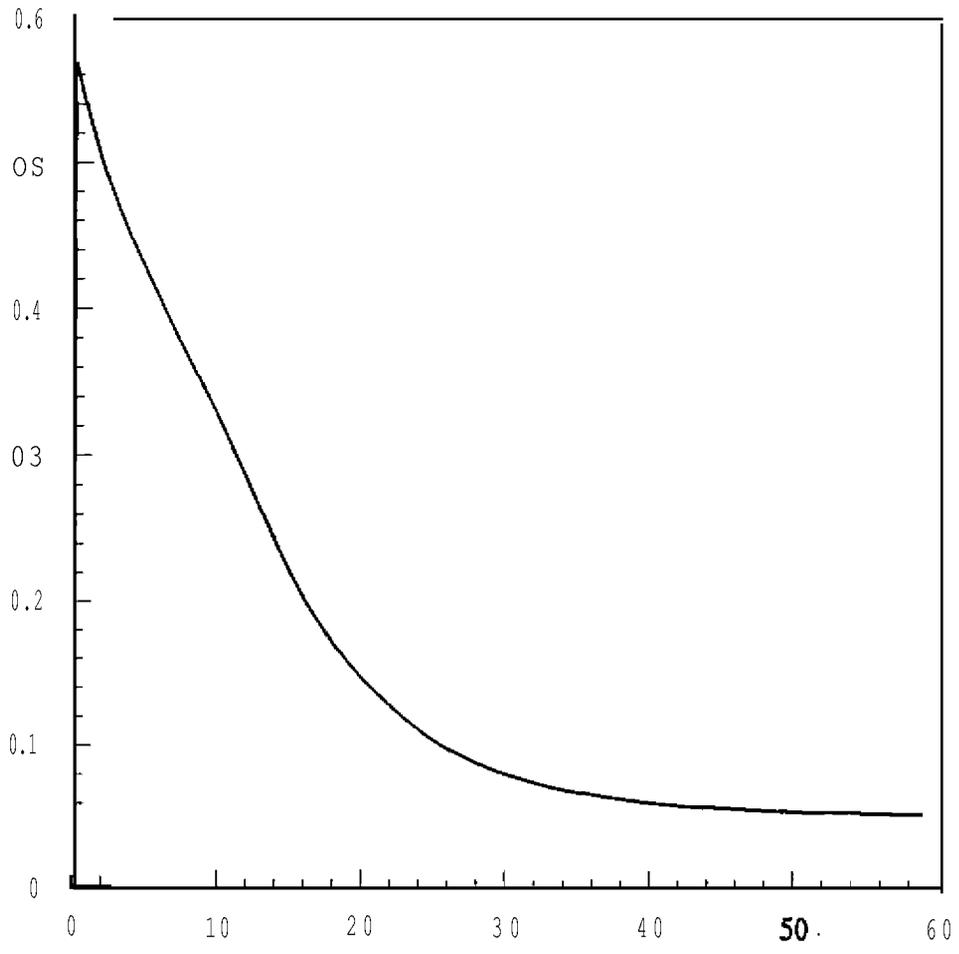


Figure 3. Example of Optimum

a) Diffusion

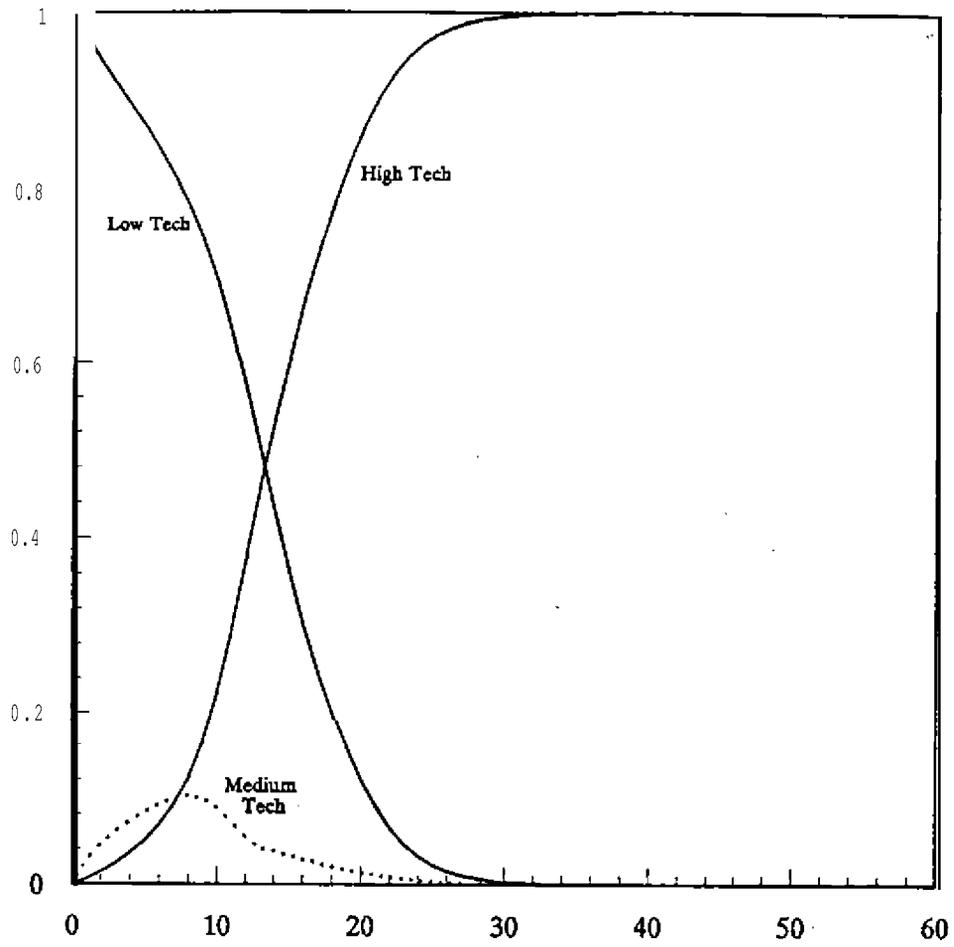


Figure 3. Example of Optimum

b) Innovation Effort (η)

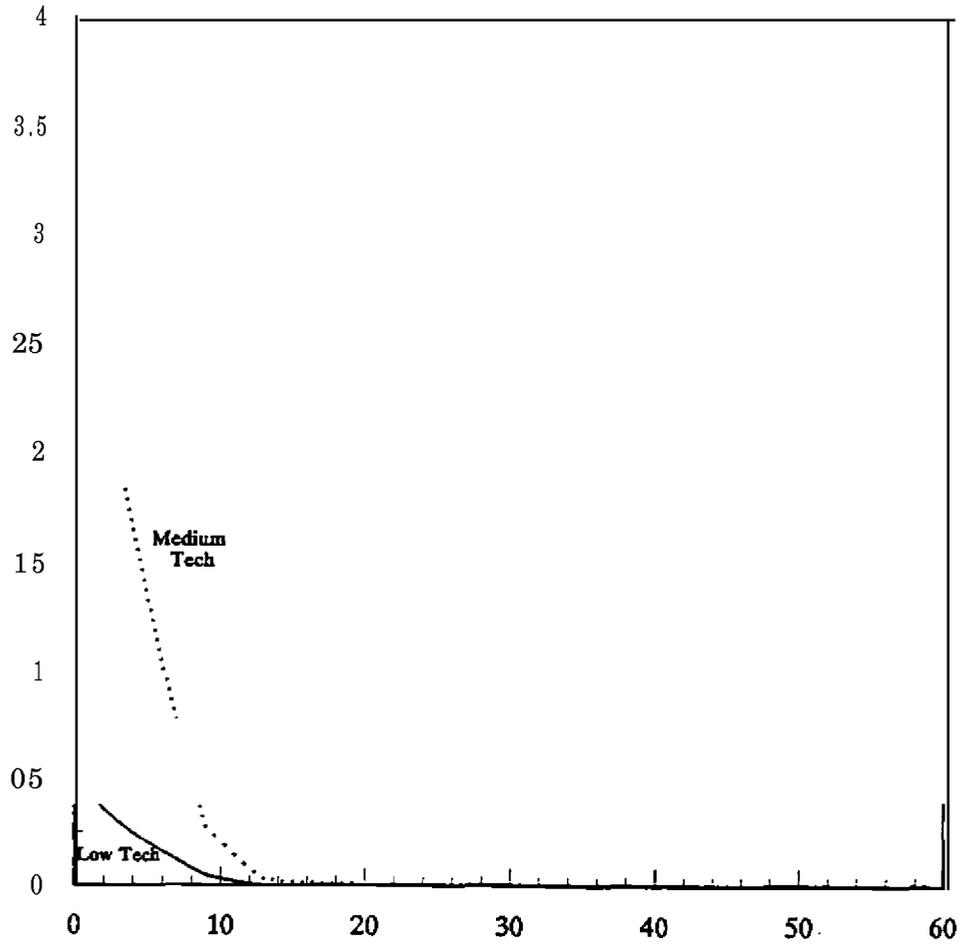


Figure 3. Example of Optimum

c) Imitation Effort (μ)

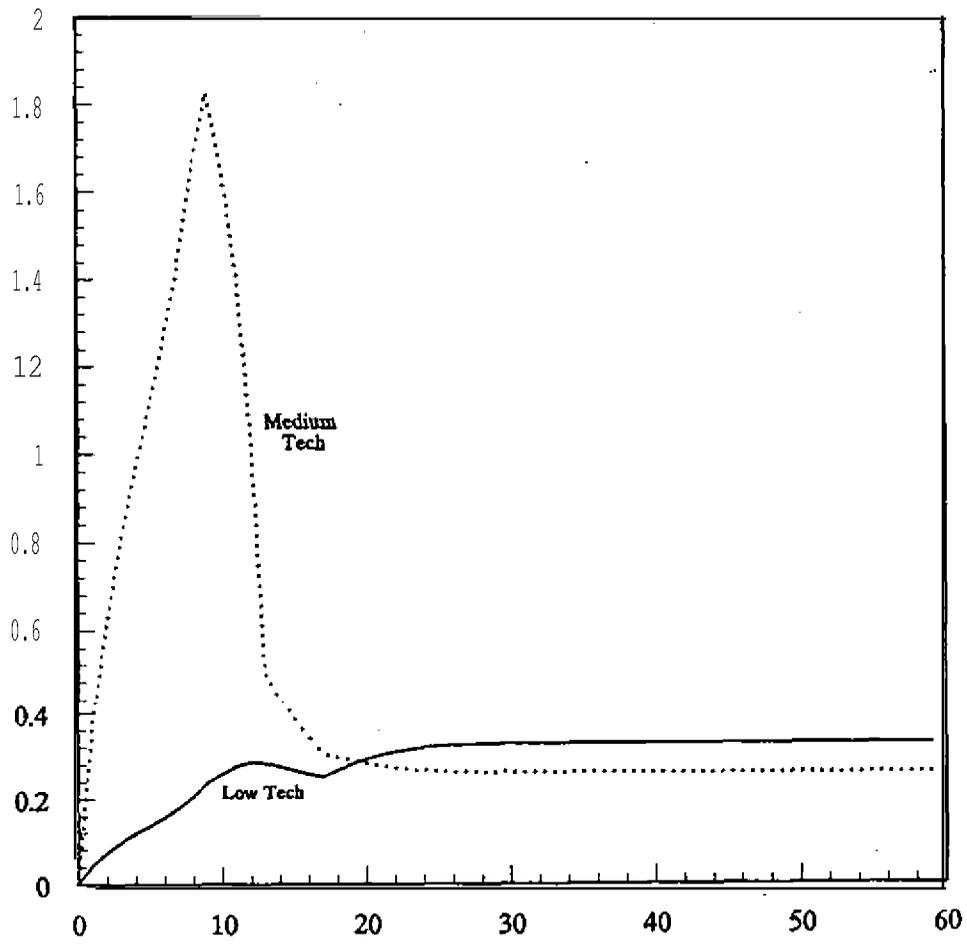


Figure 3. Example of Optimum

d) Total Output (Q) and its Variance Among Firms

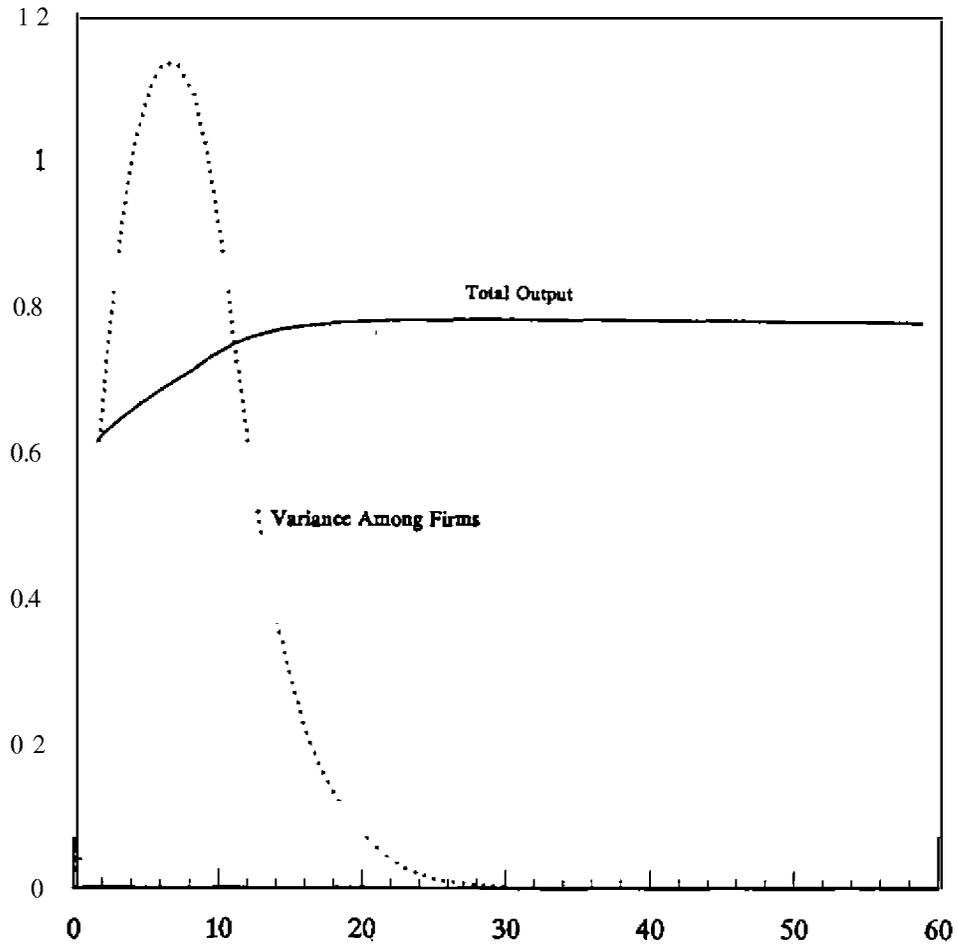


Figure 3. Example of Optimum

e) Price

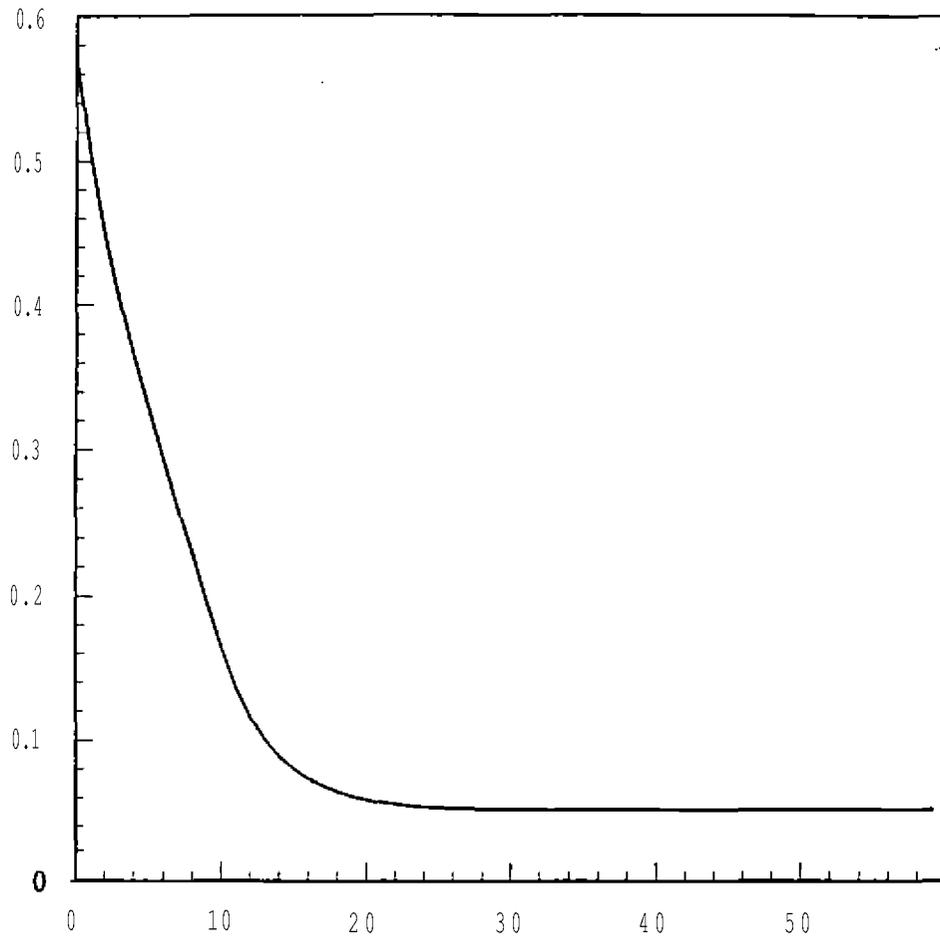


Figure 4. Diesel Locomotion in the U.S.
Railroad Industry, 1925-1966.

a) Diffusion

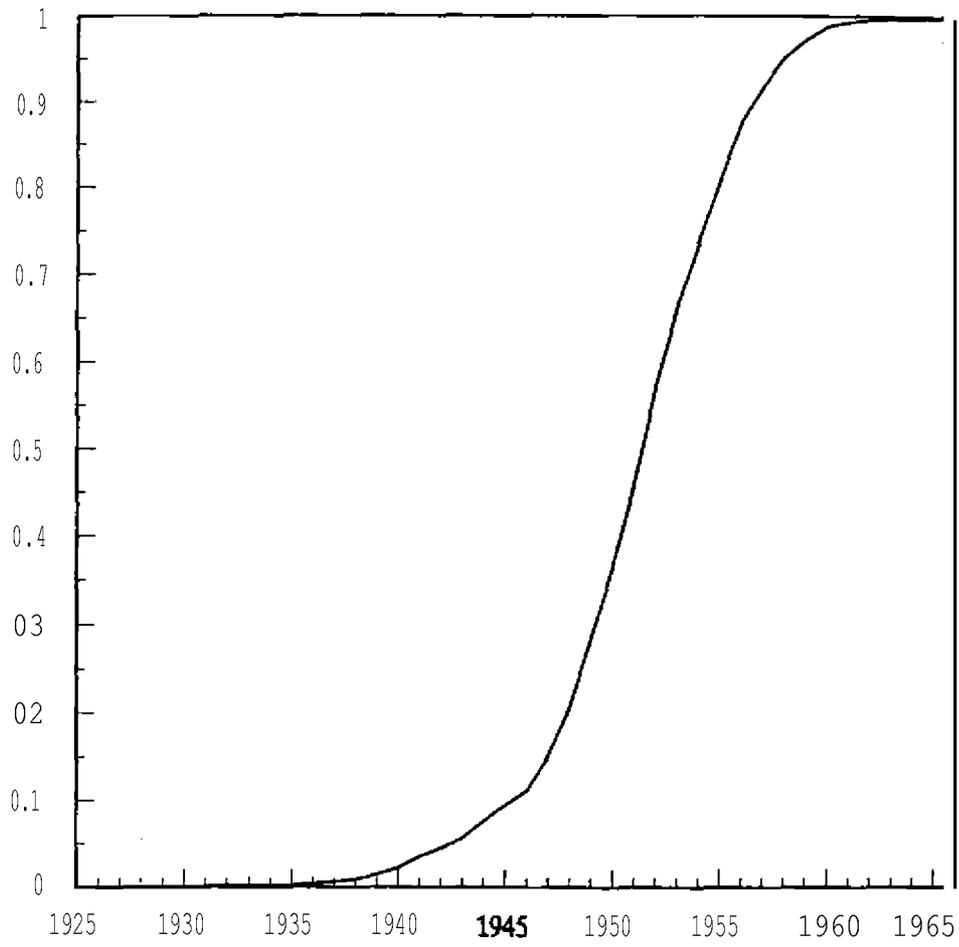


Figure 4. Diesel Locomotion in the U.S.
Railroad Industry, 1925-1966.

b) Hazard Rate

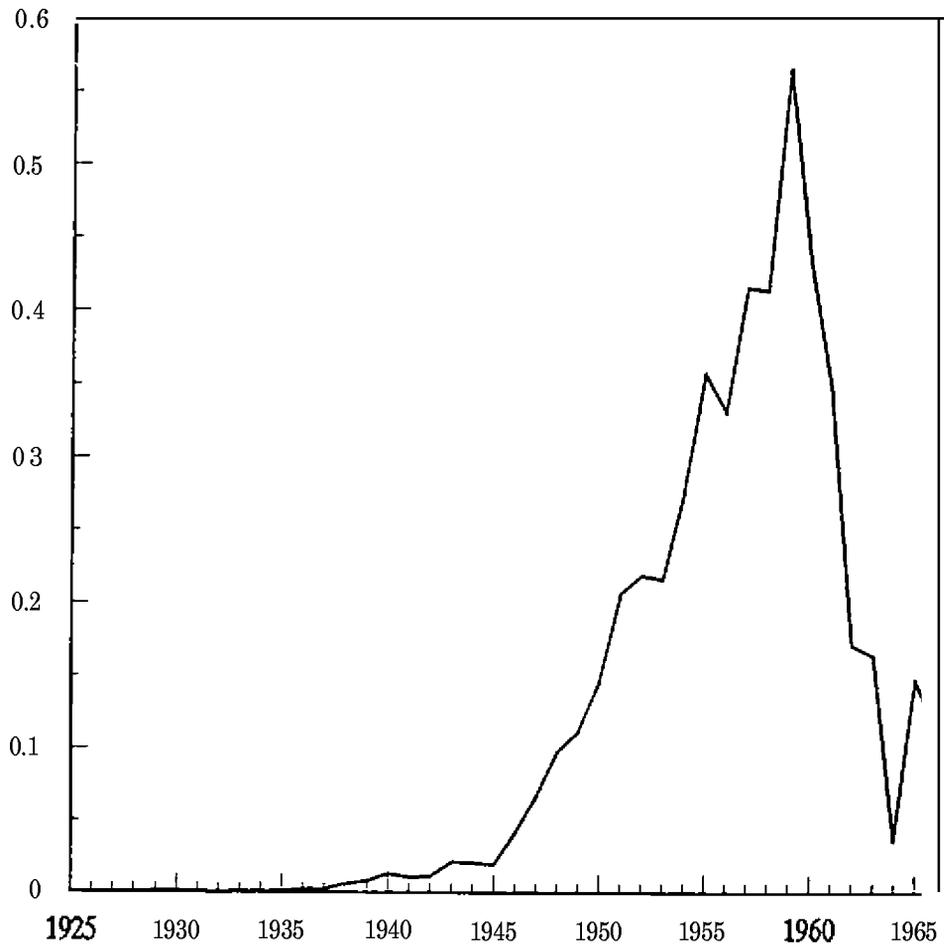


Figure 4. Diesel Locomotion in the U.S.
Railroad Industry, 1925-1966.

c) Imitation effort



<u>Number</u>	<u>Author</u>	<u>Title</u>	<u>Date</u>
4404	John H. Boyd Mark Gertler	U.S. Commercial Banking: Trends, Cycles, and Policy	7/93
440s	Alan J. Auerbach Kevin A. Hassett Stephen D. Oliner	Reassessing the Social Returns to Equipment Investment	7/93
4406	Janet Currie Duncan Thomas	Does Head Start Make a Difference?	7/93
4407	Hilary Williamson Hoynes	Welfare Transfers to Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP	7/93
4408	Charles W. Calomiris	Corporate-Finance Benefits from Universal Banking: Germany and the United States, 1870-1914	7/93
4409	Hope Corman Stephen Chaikind	The Effect of Low Birthweight on the Health, Behavior, and School Performance of School-Aged children	7/93
4410	Allan W. Kleidon Ingrid M. Werner	Round-the-clock Trading: Evidence from U.K. Cross-Listed Securities	7/93
4411	Andrew Atkeson Patrick J. Kehoe	Social Insurance and Transition	7/93
4412	Andrew B. Abel Janice c. Eberly	An Exact Solution for the Investment and Market Value of a Fii Facing Uncertainty, Adjustment Costs, and Irreversibility	7/93
4413	Steven G. Allen Robert L. Clark Ann A. McDermed	Post-Retirement Increases in Pensions in the 1980s: Did Plan Finances Matter?	8/93
4414	Edward B. Montgomery	Patterns in Regional Labor Market Adjustment: The United States vs. Japan	8/93
4415	Michael D. Bordo Charles Evans	Labor Productivity During the Great Depression	8/93
4416	Steven N. Kaplan	Top Executives, Turnover and Firm Performance in Germany	8/93
4417	Dani Rodrik	Trade and Industrial Policy Reform In Developing Countries: A Review of Recent Theory and Evidence	8/93
4418	Don Fullerton Seng-Su Tsang	Environmental Costs Paid by the Polluter or the Beneficiary? Tbe Car of CERCLA and Superfund	8/93

<u>Number</u>	<u>Author</u>	<u>Title</u>	<u>Date</u>
4419	Dani Rodrik	Trade Liberalization in Disinflation	8/93
4420	Herschel I. Grossman Murat Iyigun	The Profitability of Colonialism	8/93
4421	Takeo Hoshi Anil Kashyap David Scharfstein	The Choice Between Public and Private Debt: An Analysis of Post-Deregulation Corporate Financing in Japan	8/93
4422	Michael Bruno	Inflation and Growth in an Integrated Approach	8/93
4423	M. Ishaq Nadiri	Innovations and Technological Spillovers	8/93
4424	Olivia S. Mitchell	Retirement Systems in Developed and Developing Countries: Institutional Features, Economic Effects, and Lessons for Economies in Transition	8/93
4425	Graciela Chichilnisky Geoffrey Heal	Who Should Abate Carbon Emissions? An International Viewpoint	8/93
4426	Nancy L. Stokey Sergio Rebelo	Growth Effects of Flat-Rate Taxes	8/93
4427	Canice Prendergast Robert H. Topel	Favoritism in Organizations	8/93
4428	James Heckman	Assessing Clinton's Program on Job Training, Workfare, and Education in the workplace	8/93
4429	Rebecca M. Blank Patricia Ruggles	When Do Women Use AFDC & Food Stamps? The Dynamics of Eligibility vs. Participation	8/93
4430	Andrea Beltratti Graciela Chichilnisky Geoffrey Heal	Sustainable Growth and the Green Golden Rule	8/93
4431	Scott Boggess John Bound	Did Criminal Activity Increase During the 1980s? Comparisons Across Data Sources	8/93
4432	Janet Currie Lucia Nixon Nancy Cole	Restrictions on Medicaid Funding of Abortion: Effects on Pregnancy Resolutions and Birth Weight	8/93
4433	Magnus Blomström Edward N. Wolff	Growth in a Dual Economy	8/93
4434	James M. Poterba Lawrence H. Summers	Unemployment Benefits, Labor Market Transitions, and Spurious Flows: A Multinomial Logit Model with Errors in Classification	8/93

<u>Number</u>	<u>Author</u>	<u>Title</u>	<u>Date</u>
4435	Jonathan Gruber James M. Poterba	Tax Incentives and the Decision to Purchase Health Insurance : Evidence from the Self-Employed	8/93
4436	Magnus Blomström Robert E. Lipsey Mario Zejan	Is Fixed Investment the Key to Economic Growth ?	8/93
4437	Richard H. Clarida	Permanent Income, Import Prices, and the Demand for Imported Consumer Durables : A Structural Econometric Investigation	8/93
4438	José De Gregorio Alberto Giovannini Holger C. Wolf	International Evidence on Tradables and Nontradables Inflation	8/93
4439	Robert E. Hall N. Gregory Mankiw	Nominal Income Targeting	8/93
4440	Patric H. Hendershott William R Schultz	Equity and Nonequity Determinants of FHA Single-Family Mortgage Foreclosures in the 1980s	8/93
4441	Boyan Jovanovic Glenn M. MacDonald	The Life-Cycle of a Competitive Industry	8/93
4442	Richard H. Clarida Mark P. Taylor	The Term Structure of Forward Exchange Premia and the Forecastability of Spot Exchange Rates: Correcting the Errors	8/93
4443	V.V. Chari Lawrence J. Christiano Patrick I. Kehoe	Optimality of the Friedman Rule in Economics With Distorting Taxes	8/93
4444	David T. Coe Elhanan Helpman	International R&D Spillovers	8/93
4445	Barry Eichengreen Douglas A. Irwin	Trade Blocs, Currency Blocs and the Disintegration of World Trade in the 1930s	8/93
4446	Bernard Dumas	Partial- vs. General-Equilibrium Models of the International Capital Market	9/93
4447	Orley Ashenfelter David Bloom	Lawyers as Agents of the Devil in a Prisoner's Dilemma Game	9/93
4448	Allan Drazen Paul R. Mason	Credibility of Policies versus Credibility of Policymakers	9/93
4449	Bennett T. McCallum	Specification and Analysis of a Monetary Policy Rule for Japan	9/93

<u>Number</u>	<u>Author</u>	<u>Title</u>	<u>Date</u>
4450	Steven C. Bourassa Donald R. Haurin R. Jean Haurin Patric H. Hendershott	Independent Living and Homeownership: An Analysis of Australian Youth	9/93
4451	Dani Rodrick	Do Low-Income Countries Have a High-Wage option?	9/93
4452	Paul M. Romer	New Goods, Old Theory, and the Welfare Costs of Trade Restrictions	9/93
4453	Orazio P. Attanasio Martin Browning	Consumption over the Life Cycle and over the Business Cycle	9/93
4454	Orazio P. Attanasio	A Cohort Analysis of Saving Behavior by U.S. Households	9/93
4455	Boyan Jovanovic Saul Lach	Diffusion Lags and Aggregate Fluctuations	9/93
4456	Raquel Fernandez Richard Rogerson	Zoning and the Political Economy of Local Redistribution	9/93
4457	Louis Kaplow Steven Shavell	The Efficiency of the Legal System versus the Income Tax in Redistributing Income	9/93
4458	Bernard Dumas L. Peter Jennergren Bertil Näsland	Realignment Risk and Currency Option Pricing in Target Zones	9/93
4459	Bernard Dumas Bruno Solnik	The World Price of Foreign Exchange Risk	9/93
4460	Martin Gaynor Gerard F. Anderson	Uncertain Demand, the Structure of Hospital Costs, and the Cost of Empty Hospital Beds	9/93
4461	Kevin Lang William T. Dickens	Bilateral Search as an Explanation for Labor Market segmentation and Other Anomalies	9/93
4462	Michael D. Hurd	The Effect of Labor Market Rigidities on the Labor Force Behavior of Older Workers	9/93
4463	Boyan Jovanovic Glenn MacDonald	Competitive Diffusion	9/93

Copies of the above working papers can be obtained by sending \$5.00 per copy (plus \$10.00/order for postage and handling for all locations outside the continental U.S.) to Working Papers, NBER, 1050 Massachusetts Avenue, Cambridge, MA 02138. Advance payment is required on all orders. Please make checks payable to the National Bureau of Economic Research