Long Waves and Short Waves: Growth Through Intensive and Extensive Search
Author(s): Boyan Jovanovic and Rafael Rob
Source: *Econometrica*, Vol. 58, No. 6 (Nov., 1990), pp. 1391-1409
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/2938321
Accessed: 07/12/2010 12:54

Econometrica, Vol. 58, No. 6 (November, 1990), 1391–1409

LONG WAVES AND SHORT WAVES: GROWTH THROUGH INTENSIVE AND EXTENSIVE SEARCH

BY BOYAN JOVANOVIC AND RAFAEL ROB[1]

This paper endogenizes the frequency of major discoveries and the extent of their refinement. Four axioms deliver a one-parameter family of beliefs that guide exploratory effort. Such effort trades off the prospect of major new discovery against the chance of successfully refining discoveries made in the past. The only other parameter is the cost of making new discoveries relative to the cost of refining old ones. The paper derives time-series properties of inventive activity as they relate to the two parameters, and it discusses several specific inventions and their subsequent refinement. In doing so, the paper arguably enhances our understanding of the process of discovery.

1. INTRODUCTION

SCHUMPETER (1939) THOUGHT THAT MAJOR DISCOVERIES would be followed by waves of imitation. The unit of analysis was, for him, an entire economy; it would, in his view, be subjected to cycles in activity as long as 50 years. One can also view technical progress in a sector or industry through Schumpeter's lens; the same logic applies: invention causes a wave of imitations.

For Kuznets (1940) these ideas had intuitive appeal, but he thought that they were of little use in understanding business cycles, or even waves of activity in particular industries. His basic criticism was (a) that in reality, business cycles seemed to occur fairly regularly, and yet Schumpeter gave no reason why major inventions would be bunched at regular intervals, and (b) that Schumpeter did not really explain how the length and amplitude of cycles was related to the underlying characteristics of the economy or industry. In other words, what Kuznets found lacking was a theory with some *quantitative* predictions about the time-series properties of aggregates, or, for that matter, about the time-series pattern of productivity growth in an industry or sector.[2]

The present paper tries to be explicit and quantitative about this link: the link between the nature of discovery and imitation on the one hand, and the length and amplitude of cycles in business activity on the other. At present, however, the model is geared less towards explaining waves of *general* business activity (i.e., business cycles), then it is towards understanding the growth of productivity in industries or more narrowly defined sectors of the economy.

The argument goes as follows. A set of axioms is imposed on the formation of beliefs about technological possibilities. These axioms lead to a one-parameter (σ) family of beliefs. This parameter also represents technological opportunity.

[1] We thank the C. V. Starr Center for Applied Economics for technical and financial assistance. The second author wishes to acknowledge the financial support of NSF under Grant No. SES 8821233. We also thank Yaw Nyarko for useful remarks at an early stage, Ray Atje for capable assistance, Robin Cowan, and especially two referees for helpful comments.

[2] "The core of the difficulty [with Schumpeter's work] seems to lie in the failure to forge the necessary links between the primary factors and concepts (entrepreneur, innovation, equilibrium line), and the observable cyclical fluctuations in economic activity." (Kuznets (1940, p. 270).)

An appealing feature of the resulting beliefs is that they are essentially the same as the estimates made routinely in practice in geostatistics, meteorology, and elsewhere (Cressie (1986), Cressie and Horton (1987)). In these disciplines little is known about functional form, and prior notions similar to our axioms appear to be used in practice. The only other parameter of the model, c, represents the cost of engaging in the activity of discovery or invention relative to the cost of engaging in implementation, refinement, or imitation.

The time-series properties of business activity are then derived in terms of σ and c. A sufficiently high σ/c ratio is needed to get long-run productivity growth in the first place. If σ/c is in the intermediate region (see Figure 1), we get Schumpeter's cycles in activity. But as σ/c gets higher still, cycles disappear altogether because new inventions then appear in a steady stream—too steady to produce cycles. When cycles do occur, the industry oscillates between epochs when invention is the dominant activity, and epochs when refinement of past inventions dominates. Growth turns out to be higher during times when technologies are being refined. The reason is that the expected payoff to discovery is constant, whereas when agents choose to refine *past* inventions, they do so because refinement promises an unusually high payoff. Productivity grows slowly in those periods in which no good new discoveries or refinements are made; it tends to grow the fastest during periods in which past discoveries are refined.

Examples of earlier search-theoretic work on the growth of knowledge are Nelson (1982), Telser (1982), and Jovanovic and Rob (1989). One way to view the present contribution is that it formalizes the distinction between extensive and intensive search, a distinction made by Rosenberg (1972) among others. Extensive search seeks major breakthroughs, while intensive search attempts to refine such breakthroughs. But this distinction is useful only if it helps us understand how knowledge has grown in some fairly specific contexts. So, after presenting the model in the next two sections, we go on in Section 4 to discuss aspects of historical experience that our theory helps understand. The fifth and

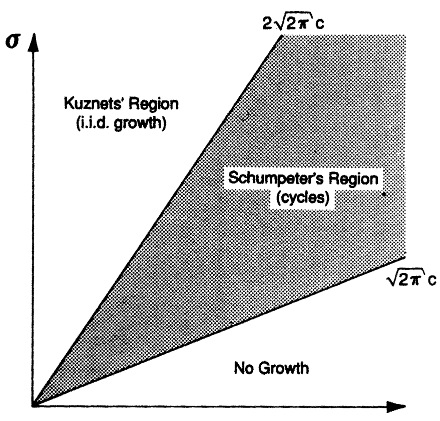

FIGURE 1.—Growth and the two parameters.

final section of the paper discusses problems and extensions, and offers some concluding remarks.

2. TECHNOLOGIES AND BELIEFS

This section focuses on the relation between technology-types and the output that they yield. The first subsection formally defines this relation. The second subsection states four axioms that this relation obeys, and that agents' prior beliefs will recognize; these axioms lead agents to a unique prior over functions relating technology-types and output. Finally, the third subsection describes the choices that agents face.

2A. *Technologies*

There is a single output and a variety of technologies which can be used to produce it. Each technique is represented by an infinite-dimensional vector, $x \equiv (x_1, x_2 \dots)$, where $0 \leqslant x_i \leqslant 1$. We are assuming, thus, a countable infinity of technology-types, and a continuum of each type. The output associated with x is $z(x)$. $z(\cdot)$ is *not* an input-output relationship. Rather, x is a *method* of production, or more precisely, a combination of such methods, whereas $z(x)$ is its *net* (of input costs) productivity. In particular, z can assume negative values.

Some concrete examples will help fix ideas. Suppose that dimension k refers to drilling oil at location k. Then $x_k \in [0,1]$ could be the depth at which drilling takes place, where 1 represents the maximum depth at which drilling is feasible. Alternatively, x_k can be thought of as the angle at which an arrow is shot. Then $x_k = 0$ could represent a direct aim (which invariably produces a miss at a significant distance), and $x = 1$ an aim vertically up in the air. Dimension k will be referred to as a technology-type, say for drilling oil or for archery. We are thus assuming a countable infinity of technology-types.

The "universe of techniques," $[0,1]^\infty$, is a priori known. But what they yield, i.e., the function $z(\cdot)$, is not. That is, $z(\cdot)$ is a random function. Certain restrictions, pertaining to variations in a single component of x, will be imposed on $z(\cdot)$. Taken together they will lead to a prior measure over the outputs of all techniques.

Let $\Delta_n \equiv \{x \in [0,1]^\infty | x = (x_1, \dots, x_n, 0, 0, \dots)\}$, $\Delta = \bigcup_{n=1}^{\infty} \Delta_n$, and for any $x \in [0,1]^\infty$, any positive integer k, and any y_k, let $(x | y_k) \equiv (x_1, \dots, x_{k-1}, y_k, x_{k+1}, \dots)$.

2B. *Beliefs about Technological Possibilities*

Any parametric family of z's along with a prior over its parameters would imply a prior over the functions z. Such a prior typically assigns measure zero to a lot of functions $z(\cdot)$. We wish to *derive* the prior beliefs from postulates (deemed to be held by the agents in the model) about the nature of production. It is well known that in infinite-dimensional spaces there is no unique way to express complete ignorance.[3] To get useful results, something must be assumed. On the other hand, beliefs must include a large enough collection of $z(\cdot)$ functions so that things that do not seem too unreasonable a priori are included. Our approach is similar to methods followed in geostatistics[4] in which predicted distributions of reserves of oil or ore at unexplored locations are formed roughly in the same way that our agents form beliefs about technologies they have not yet tried out. Similar methods are used in hydrology and meteorology.

The following axioms are imposed on beliefs.

ASSUMPTION 1 (Continuity): *$z(\cdot)$ is continuous in each variable separately. Thus, techniques are given locational context (setting a dial on a machine or drilling for oil in a certain location, for instance) and a slight change in x_k is assumed not to produce a dramatic change in output.*

ASSUMPTION 2 (Zero Drift): *For each $x \in \Delta$, each k, and each x_k' such that $x_k' > x_k$, $E\{z(x|x_k')|z(x)=z\}=z$. This axiom expresses complete ignorance about whether a new technology, or the further development of an existing technology (in the direction of a larger x_k), will raise output or reduce it.*

ASSUMPTION 3 (Constant Proportional Uncertainty): *$Var\{z(x|x_k')|z(x)=z\} = \sigma^2(x_k'-x_k)z^2$, where $\sigma>0$ is a given constant and $x_k'>x_k$. This makes the standard deviation of the output resulting from the trial (in dimension k) proportional to z, and to $(x_k'-x_k)^{1/2}$. The proportionality to z implies that as z grows, more will be at stake as one experiments with a new technology. This captures the well-known argument that returns to information are proportional to the operating scale at which the information is used (Wilson (1975)). The proportionality of the variance to $(x_k'-x_k)$ means that in each dimension, sampling far away from the previously-known technology x_k leads to greater variance. The fact that this variance is linear in $x_k'-x_k$ is just a matter of choosing units of x appropriately.*

ASSUMPTION 4 (Independent Increments): *Let $x_k' < x_k < x_k''$. Then $z(x|x_k') - z(x)$ and $z(x|x_k'') - z(x)$ are independent. This axiom expresses another aspect of maximum ignorance. An increase, say, in output as one moves from x_k' to x_k contains no information on what will happen to output if we should experiment with x_k''.*

REMARK 1: We assume throughout that σ is known by the agents. If it were unknown, precise inference about it would be made fairly quickly (say within 50 periods), so that our model captures whatever takes place following these initial periods.

[3] Even on the line, there is a large collection of measures that assign zero measure to each point.

[4] Geostatistics is the method used to analyze reserves of ore and oil in the ground, and to predict reserves at hypothetical locations given observations at certain other locations. The analysis there is usually in R^2 or R^3 but the concepts readily generalize to higher dimensions. See Cressie (1986), and Cressie and Horton (1987).

REMARK 2: All but the first assumption are commonly made in geostatistics. The continuity assumption would not be appropriate there if the geographical structure were riddled with faults leading to sudden jumps and discontinuities underground. Otherwise it too is reasonable in the physical context. We need it to fully nail down beliefs as represented in (2.1) below. If, instead, discontinuous z's were deemed reasonable, we would be dealing with jump processes, and further assumptions would be needed to nail down beliefs.

LEMMA 1 (Billingsley (1968, p. 154)): *The above four assumptions imply that for each k, $[z(x)-z(x|0_k)]/z(x|0_k)$ is Brownian Motion with incremental variance σ^2; thus the percentage increase in output follows Brownian Motion in each technological dimension.*

COROLLARY: *The explicit unique representation of $z(\cdot)$ is*

$$z(x)=\prod_{k=1}^{\infty}\left[1+\sigma W_k(x_k)\right], \quad x\in\Delta, \tag{2.1}$$

where $(W_k(\cdot))_{k=1}^{\infty}$ is a sequence of sample paths of Brownian motions with $W_k(0)=0$, all k, and where $\sigma>0$.

PROOF: From the Lemma, we have $z(x)=z(x|0_k)[1+\sigma W_k(x_k)]$ for all $k=1,2,\dots$ and for all x. But $z(x|0_k)=z(x|0_k,0_j)][1+\sigma W_j(x_j)]$ for all $j\neq k$. Since $x\in\Delta$, we can, through a finite number of substitutions for z, reach equation (2.1) as the unique representation. *Q.E.D.*

REMARK 1: Equation (2.1) says nothing about possible forms of dependence amongst the W_k (e.g., symmetric, or geometrically declining in k, etc.). Such dependence allows for a sort of transfer of knowledge (i.e., inferences) across technologies. While we shall comment later on the likely consequences of such dependence, our formal analysis will assume that the W_k are *mutually independent*.

REMARK 2: The ordering of the possibilities in the kth dimension is in the direction of increased subjective uncertainty. Technique k yields zero for sure if x_k is set at zero. The larger x_k, the larger is the uncertainty about the outcome.

REMARK 3: The parameter σ is thought of as measuring technological opportunity. Since σ is not indexed by k, every technology is ex ante equally promising. This is a consequence of the third axiom.

REMARK 4: Although technological discoveries interact because (2.1) is of a multiplicative form, neither current output nor (as we shall soon see) the prospects for future discovery depend on the order in which past discoveries were made. This will rule out certain kinds of "path dependence" in optimal search policies.

2C. *Choices Available to Agents*

Our aim is to look into Schumpeter's assertions about invention and cycles of activity, given what we regard as reasonable assumptions about the process of discovery and refinement. We do this in the simplest possible way, by having effectively just one agent. This agent could be Robinson Crusoe who consumes his own output, but a better interpretation is that of a firm that can appropriate the fruits of its search efforts for exactly one period. Later we shall comment on the likely effects of the presence of many agents, and of longer time horizons.

There are overlapping generations of risk-neutral agents that live one period. Each generation consists of exactly one member and each such member can make exactly one search. Search means selecting a new technique x, and observing its net productivity, $z(x)$. As a result of past searches, at each point in time there is a body of empirical knowledge,[5] $H^t = \{x^1, z^1, \dots, x^t, z^t\}$, where $z^t = z(x^t)$. The history H^t is assumed to be *known* to generation t and, thus, this generation's consumption is

$$(2.2) \qquad Z_t \equiv \max(z^1, \dots, z^t).$$

We are assuming, then, that information is costlessly passed on from generation to generation. Hence, each generation can exploit the best available technique hitherto sampled. In contrast to the multi-armed bandit formulation in which the agent is forced to consume the payoff of the arm that he pulls, our model unbundles consumption from search. Empirically, this is probably realistic: One is not forced to use an unprofitable technique.

Let $w_j^* = \max_{x \in H^t} W_j(x_j)$. We assume that $x_k = 0$ is, for any k, an option that is always available to an agent. Hence, $w_j^* \geqslant 0$. Using expression (2.1) we can then rewrite (2.2) as

$$(2.3) \qquad Z_t = \prod_{j=1}^{n_t} (1 + \sigma w_j^*).$$

This is what generation $t+1$ can guarantee itself prior to search.

Two modes of search are available: intensive and extensive. Intensive search means experimentation along an *old* technological dimension. For intensive search, a technology-type which had been sampled before is selected, and experimentation is conducted with a new technique belonging to it. This search is costless. Let n_t be the number of technology-types hitherto sampled at least once. If generation $t+1$ chooses to search intensively, it must select a vector $x^t \in H^t$, a coordinate $1 \leqslant k \leqslant n_t$, and a value $x_k^t \in [0,1]$. It then observes $z^{t+1} = z(x^t | x_k^t)$. That is, it can vary x in *one dimension at a time*. For extensive search, experimentation with a different technique is done along a *new* technological dimension, i.e., using a technology type about which nothing is known. Each extensive search costs cZ_t, where $0 < c < 1$. For this type of search a

[5] Note that the output of each technique is observed exactly, without error.

vector $x^t \in H^t$ and a value $x_{n_t+1} \in [0,1]$ are chosen and $z^{t+1} = z(x^t | x_{n_t+1})$ is observed.[6]

3. ANALYSIS

The decision facing each generation is whether to sample extensively or intensively, and given the chosen mode of search, exactly which technique to sample. The payoff to each type of search will now be described in turn.

3A. *Optimal Extensive Search*

If extensive search is the chosen option, then the following theorem holds.

THEOREM 1: $x_{n_t+1} = 1$, *and the expected payoff to extensive search is*

$$Z_t(1 + \sigma/\sqrt{2\pi} - c) = \prod_{j=1}^{n_t} (1 + \sigma w_j^*)(1 + \sigma/\sqrt{2\pi} - c). \tag{3.1}$$

PROOF: Let z' denote the productivity of next period's search. When searching extensively, $z' = Z_t(1 + \sigma W_{n_t+1}(x_{n_t+1}) - c)$. Since $W_{n_t+1}(x_{n_t+1}) \sim N(0, x_{n_t+1})$ and is independent of prior history, we find

$$E\{Z_t(1 + \max(0, \sigma W_{n_t+1}) - c) | x_{n_t+1}\} = Z_t\left[1 + \sigma x_{n_t+1}^{1/2}/\sqrt{2\pi} - c\right], \tag{3.2}$$

where the above equality follows from a straightforward calculation.[7] The assertion follows. *Q.E.D.*

Since intensive search is costless, and its expected payoff is hence at least Z_t, a necessary condition for extensive search to ever be chosen is the following assumption:

ASSUMPTION 5: $\sigma/\sqrt{2\pi} \geqslant c$.

This assumption is maintained for the rest of this section.

3B. *Optimal Intensive Search*

A history, H^t, induces a partition on each of the first n_t coordinates. Sufficient statistics for the beliefs concerning the outcome of sampling within

[6] The reason for assuming a search cost proportional to Z is that in practice, R&D is a highly labor-intensive activity, so that its social cost is essentially proportional to the foregone output that the scientists and engineers engaged in R&D could otherwise be producing. Moreover, Kuznets (1962, pp. 31–35), when discussing the problems of measuring the input into inventive activity, thought that the costs *not* captured in measured R&D were even *more* weighted towards the foregone labor input—the input of individual and independent inventors. At any rate, our assumption is that this foregone-output cost is incurred each time an extensive search is made.

[7] If $\varepsilon \sim N(m, s^2)$ and $\bar{\varepsilon}$ is a constant, then $u(m, s, \bar{\varepsilon}) \equiv E \max(\varepsilon, \bar{\varepsilon}) = m + (\bar{\varepsilon} - m)F((\bar{\varepsilon} - m)/s) + (s/\sqrt{2\pi})\exp\{-(\bar{\varepsilon} - m)^2/2s^2\}$, where F is the standard normal CDF.

each interval of that partition are the values of z at its endpoints. This follows from the fourth assumption.

Three stages are involved in intensive search. Stage 1: the agent selects a coordinate k, $1 \leqslant k \leqslant n_t$; Stage 2: he selects an interval belonging to the history-induced partition of k; Stage 3: he chooses a value x_k' within that interval.

Prior to search at t, the agent can guarantee himself consumption Z_t (see expression (2.3)). If the sampled technology yields z', following an intensive search in dimension k at technology x_k' he gets

$$\max(z', Z) = \prod_{j \neq k} (1 + \sigma w_j^*)[1 + \sigma \max(W_k(x_k'), w_k^*)], \tag{3.3}$$

(since $z' = \prod_{j \neq k}(1 + \sigma w_j^*)[1 + \sigma W_k(x_k')]$). Thus, letting $\pi(H^t)$ be the expected payoff to intensive search at $t+1$, we have

$$\begin{aligned} \pi(H^t) &= \max_{1 \leqslant k \leqslant n_t} \left(\max_{x_k} E[\max(z', Z)] \right) \\ &= \max_{1 \leqslant k \leqslant n_t} \prod_{j \neq k} (1 + \sigma w_j^*) \left[1 + \sigma \max_{x_k'} E\{\max[W_k(x_k'), w_k^*]\} \right]. \end{aligned} \tag{3.4}$$

Exposition is easiest if the final, third stage of intensive search is discussed first. When sampling along the kth dimension, we are learning about $W_k(\cdot)$, because of the multiplicative separability in equation (2.1). Let $[\alpha, \beta] \subset [0, 1]$ be a subinterval in the kth dimension, with $W_k(\alpha) = W^\alpha$ and $W_k(\beta) = W^\beta$. That is, W^α and W^β are values associated with previously-experimented with technologies $x_k = \alpha$ and $x_k = \beta$. (Note that intensive sampling can never be in an interval with an unobserved endpoint, because $W_k(0) = 0$ by Lemma 1, while extensive search of k must precede intensive search of k, and *it* yields an observation of $W_k(1)$, by Theorem 1.)

Conditional on intensive sampling within $[\alpha, \beta]$, the choice of x_k induces a W_k whose distribution conditional on (W^α, α) and (W^β, β) is normal (see Billingsley (1968, p. 65) for details on the Brownian bridge) with mean

$$m = W^\alpha(\beta - x_k)/(\beta - \alpha) + W^\beta(x_k - \alpha)/(\beta - \alpha), \tag{3.5}$$

and with variance

$$s^2 = (\beta - x_k)(x_k - \alpha). \tag{3.6}$$

Thus, given the dimension k and an interval $[\alpha, \beta]$, it is evident that the maximization of (3.4) is equivalent to maximizing $u(m, s, w_k^*)$ (see footnote 7) subject to the constraints (3.5) and (3.6). Let $v(\alpha, \beta, W^\alpha, W^\beta, w_k^*)$ be the maximized value of that program; $v(\ )$ is the incremental percentage value of intensive search.

3C. *Intensive Versus Extensive Searches*

It is now time to compare the two modes of search. Comparing (3.1) and (3.4) we see that intensive search on $[\alpha, \beta]$ will take place only if

$$1 + \sigma v(\alpha, \beta, W^\alpha, W^\beta, w_k^*) > (1 + \sigma w_k^*)(1 + \sigma/\sqrt{2\pi} - c). \tag{3.7}$$

(We have eliminated the multiplicative factor $\prod_{j \neq k}(1+\sigma w_j^*)$ which is common to (3.1) and (3.4).) Note that inequality (3.7) is time invariant, so that this proves the following theorem.

THEOREM 2: *Once a new dimension (technique) is explored, none of the previous dimensions will ever be further explored.*[8]

Refinement of techniques: This completes the discussion of the third stage of intensive search. We now discuss the first two stages. Not every newly-discovered technique will be further explored or refined. Those which, upon their discovery, are developed further belong to the set

$$D \equiv \{\omega \in R \,|\, v(0,1,0,\omega,\max(0,\omega)) > (1+\sigma/\sqrt{2\pi}-c)\max(0,\omega)+1/\sqrt{2\pi}-c/\sigma\},$$

where $\omega \equiv W_{n_t+1}(1)$. Note that the agent can always guarantee himself at least $\max(0,\omega)$ from technology n_t+1.

THEOREM 3: (i) *D is nonempty if and only if $c > \sigma/2\sqrt{2\pi}$.*

(ii) *In that case, $D=[\underline{w},\overline{w}]$ where $\underline{w}<0<\overline{w}$, and $\underline{w}$ and $\overline{w}$ are the two solutions for ω to the equation*

$$v[0,1,0,\omega,\max(0,\omega)] = (1+\sigma/\sqrt{2\pi}-c)\max(0,\omega) + 1/\sqrt{2\pi}-c/\sigma. \tag{3.8}$$

PROOF: The "if" part of (i) is shown by demonstrating that $v(0,1,0,0,0) > 1/\sqrt{2\pi}-c/\sigma$. But because the optimal $x_k^{\prime}$ is then $1/2$, and equation (3.6) yields $s^2=1/4$, footnote 7 implies $v(0,1,0,0,0)=\sigma/2\sqrt{2\pi}$, and the assertion follows. The "only if" part of (i) is demonstrated by looking at the derivatives of the two sides of (3.8). By applying the envelope theorem to v, we find that

$$\partial v/\partial\omega = (1-F)x_k + I(\omega)F \tag{3.9}$$

(where $I(\omega)=1$ if $\omega \geqslant 0$, and zero otherwise). Also,

$$\partial \max(0,\omega)/\partial\omega = I(\omega). \tag{3.10}$$

Clearly, if v is not above the right-hand side of (3.8) at $\omega=0$, it cannot exceed it for any ω, because for $\omega>0$, the right-hand side of (3.9) is no greater than the right-hand side of (3.10) (which, in turn is equal to 1), and for $\omega<0$, the right-hand side of (3.9) is nonnegative while the right-hand side of (3.10) is zero. This proves the "only if" part of (i).

[8] This theorem asserts in effect that "recall" to past unexplored opportunities does not matter. Also, equation (3.7) compares single-period returns to the two modes of search. This remains the optimal way agents would compare the two options even if they lived for more than one period, so long as the gains from either type of search could not be appropriated for more than one period. (This assertion is certainly not true universally; changes in relative prices can lead agents to return to previously abandoned technologies—see the *N.Y. Times* (1988, May 18, July 27, and November 22) for examples recorded in 1988.

Turning to (ii), assume that D is nonempty. The existence of $\underline{w} < 0 < \overline{w}$ solving (3.8) will follow if we can show that (a) $\lim_{\omega \to -\infty} v = 0$ and (b) $\lim_{\omega \to \infty} [v - (1 + \sigma/\sqrt{2\pi} - c)\omega] \leqslant 0$. Now (a) follows because $\lim_{m \to -\infty} u = 0$. For (b), note that the derivative of v is no greater than 1 (see (3.9)), whereas the derivative of $(1 + \sigma/\sqrt{2\pi} - c)\omega$ is strictly greater than 1 by Assumption 5. *Q.E.D.*

Next, looking at an interval $[\alpha, \beta]$, we provide a *necessary* condition for the continuation of (intensive) search on that interval.

THEOREM 4: *In order for search to take place on an interval* $[\alpha, \beta]$, *we must have*

$$(3.11) \qquad \beta - \alpha \geqslant 2(1 - c\sqrt{2\pi}/\sigma)(1 + \sigma w^*),$$

where w^* *is the maximal sampled* W *along the dimension to which the interval* $[\alpha, \beta]$ *belongs.*

PROOF: The incremental value of intensive search on $[\alpha, \beta]$ is given by $v[\alpha, \beta, W^\alpha, W^\beta, w^*]$, which cannot exceed $v(\alpha, \beta, w^*, w^*, w^*)$ (since v is increasing in W^α and W^β and since $w^* \geqslant \max(W^\alpha, W^\beta)$). Furthermore, when $W^\alpha = W^\beta = w^*$, the optimal choice for x_k' is $(\alpha + \beta)/2$ which by (3.5), (3.6), and footnote 7 implies

$$v(\alpha, \beta, w^*, w^*, w^*) = w^* + (\beta - \alpha)/2\sqrt{2\pi}.$$

On the other hand, the incremental value of extensive search is

$$w^*(1 + \sigma/\sqrt{2\pi} - c) + 1/\sqrt{2\pi} - c/\sigma.$$

Hence, an intensive search on $[\alpha, \beta]$ is preferred to an extensive search *only if*

$$w^* + (\beta - \alpha)/2\sqrt{2\pi} > w^*(1 + \sigma/\sqrt{2\pi} - c) + 1/\sqrt{2\pi} - c/\sigma.$$

But this, by a slight rearrangement, is equivalent to (3.11). *Q.E.D.*

In particular, setting $w^* = 0$ (which by the assumption preceding equation (2.3) is smaller than the true w^*), we get a uniform lower bound on the length of $[\alpha, \beta]$:

$$(3.12) \qquad \beta - \alpha \geqslant 2(1 - c\sqrt{2\pi}/\sigma).$$

A corollary of Theorem 3 concerns T, which we define as the (random) duration of intensive search. The largest number of times that one can sample within an interval of unit length without sampling an interval shorter than Δ is Δ^{-1} times. Therefore, taking the inverse of (3.12) yields the upper bound on the duration of Schumpeter's cycles:

THEOREM 5: $T \leqslant \sigma/2(\sigma - c\sqrt{2\pi})$, *w.p.*1.

3D. *Two-Period Generations*

This subsection will consider the case where an innovator can appropriate the value of his innovation for two periods. In that case, the preferred mode of search will be affected, of course, not just by its immediate payoff but also by the expected value of the subsequent search. Our purpose here is to compare the decision made by a one-period firm and a two-period firm at a specific point along the evolutionary path of the industry, namely, at a point where the one-period firm is *indifferent* between the two modes of search. It will be shown that under these circumstances the two-period firm prefers to do *extensive* search. This result lends support to the idea that firms guided by longer-term considerations will tend to engage in more venturesome research projects.[9]

THEOREM 6: *Assume that*

$$1+\sigma v\left(\alpha, \beta, W^{\alpha}, W^{\beta}, w_k^*\right) \leqslant\left(1+\sigma w_k^*\right)\left(1+\sigma / \sqrt{2 \pi}-c\right), \tag{3.13}$$

for all subintervals which are indexed by a given history, and assume that an equality holds for at least one such subinterval. Then the two-period payoff under initial extensive search exceeds the two-period payoff under initial intensive search.

PROOF: From the discussion leading to equation (3.7) above, it is clear that the first-period payoff, i.e., Z_t is identical under the two modes of search. Hence, it only remains to compare the second period payoffs. We start out by proving the following.

CLAIM: If intensive search is initially undertaken, then the next search will *necessarily* be *extensive* (this is a stronger form of the "no going back" property that Theorem 2 asserts).

This claim certainly holds if the knowledge gained as a result of the initial (intensive) search does not represent an improvement upon the previously best-known method of production along the same technological dimension, i.e., if $W_k(x_k^*) \leqslant w_k^*$. Concerning the case where $W_k(x_k^*) > w_k^*$, it is clear that an intensive search in the second period could possibly take place only in (α, x_k^*) or in (x_k^*, β). We now rule out the possibility of a profitable intensive search in (α, x_k^*); the proof for (x_k^*, β) is perfectly analogous.

Consider the derivatives with respect to w_k^* of both sides of (3.13):

$$\partial / \partial w_k^*\left\{\left(1+\sigma w_k^*\right)\left(1+\sigma / \sqrt{2 \pi}-c\right)\right\}=\sigma\left(1+\sigma / \sqrt{2 \pi}-c\right) \geqslant \sigma,$$

using Assumption (B.1); and

$$\begin{aligned} & \partial / \partial w_k^*\left\{\sigma v\left(\alpha, x_k^*, W^{\alpha}, w_k^*, w_k^*\right)\right\} \\ & \quad=\sigma\left\{F\left(\left(w_k^*-m^*\right) / \sigma^*\right)\right. \\ & \quad\quad\left.+\left[1-F\left(\left(w_k^*-m^*\right) / \sigma^*\right)\right]\left(y_k^*-\alpha\right) /\left(x_k^*-\alpha\right)\right\}, \end{aligned}$$

[9] A complete analysis of the two period case is rather involved because of both the complexity of the "state space" (which comprises all possible histories of the search process) and the mixed nature of each firm's choice, consisting of n discrete and a continuous decision. An attempt to study this problem led to a highly cumbersome formulation, and we were not able to consummate the analysis.

where y_k^* is the payoff maximizing search in (σ, x_k^*) and where $m^* = m(y_k^*)$, $\sigma^* = s(y_k^*)$. In computing the second derivative, we had used the envelope theorem, recalling that v is the maximized value of $u(m, s, w_k^*)$ (consult, again, the analysis leading to equation (3.7)).

Since (3.13) held with equality prior to the initial period search, the above computation shows that it will hold with strict inequality after that search. Hence, the second period search must be extensive, and the claim is proved.

Returning to the proof of the theorem, we now know that the second period payoff under intensive search is $Z_{t+1}(1 + \sigma/\sqrt{2\pi} - c)$. On the other hand, Theorem 3 shows that the second period return to extensive search is *at least* $Z_{t+1}(1 + \sigma/\sqrt{2\pi} - c)$, and is actually higher whenever the outcome of the initial period (extensive) search is such that $\underline{w} < W_{n_t+1}(1) < \overline{w}$. Furthermore, $\underline{w} < W_{n_t+1}(1) < \overline{w}$ is a positive probability event. *Q.E.D.*

This subsection has, strictly speaking, analyzed the change in the optimal search policy when instead of one-period, there are two-period non-overlapping generations. These generations have, in effect, monopoly power to perform two search decisions in a row. This comparison is meant to be only suggestive of the likely effects of longer-term patent protection on the nature of inventive activity that firms undertake.

4. EMPIRICAL IMPLICATIONS AND EVIDENCE

This section's aim is to convert the theorems proved in the last section into propositions about observables. This will now be done in two ways. The first subsection will discuss some fairly general features that one might look for in the time-series of output and inventive intensity. Then, the second subsection will discuss some specific major innovations and their subsequent refinement.

4A. *General Time-Series Properties*

One aim of this paper was to relate, qualitatively and quantitatively, Schumpeter's waves to the nature of the technology and to the way agents learn about it. How long is each wave of activity likely to last, and when will waves exist at all? Kuznets wondered if the occurrence of inventions would not be too regular to produce waves. How do the parameters of the model bear on these opposing views?

Periods of extensive search are usefully thought of as periods of invention. An invention of technique k that has x_k falling in the interval $[\underline{w}, \overline{w}]$ will be followed up by further refinement of that technique, which one might associate with more minor innovations, or imitation. We associate with Schumpeter the outcome that the industry oscillates between the state of the extensive search (E) and the state of intensive search (I). Each uninterrupted spell in state I is a "wave" of activity sparked by the discovery of a new technique. Not all discoveries lead to such waves: Only those techniques, k, with $W_k(1) \in D$ will

	E	I
E	$F(\underline{w}/\sigma)+1-F(\overline{w}/\sigma)$	$F(\overline{w}/\sigma)-F(\underline{w}/\sigma)$
I	$1-Q_T$	Q_T

FIGURE 2

lead to transitions into I. Under this interpretation of Schumpeter's long waves, *such waves will exist if and only if* $c>\sigma/2\sqrt{2\pi}$ (see Theorem 2(i)); by Theorem 5, *they can last at most* $\sigma/2(\sigma-c\sqrt{2\pi})$ periods. Figure 1 summarizes the parametric configuration necessary to produce long waves of activity.

In the figure, the region in which there is no growth should be dismissed as empirically irrelevant, at least at present, when positive growth on average is almost universal in industries. Since both extensive search (which produces the spark) and intensive search (which is defined as the long wave) are necessary for long waves to exist, the (c,σ) pairs must be in the shaded region, Schumpeter's region, whose shape is based on the inequality in Assumption 5, and on Theorem 3(i).

If $c\to\infty$, the set D (defined just prior to Theorem 3) becomes the entire line, so that the (1, 1) cell of the above matrix becomes zero. The industry will never be in state E, and we are in the southeast region of Figure 2. Moreover, if the industry is in I, any growth that takes place will be a short-run phenomenon, because along each technological dimension, the sample paths are bounded with probability one. On the other hand, if the parameters are such that D is empty, the economy will always stay in E, and we will get serially uncorrelated, i.i.d. long-run growth rates, with mean given by $\sigma/\sqrt{2\pi}-c$. The northwest region may thus be termed Kuznets' region, because new discoveries are then occurring too frequently to permit waves to exist in between.

When $\sigma>c\sqrt{2\pi}$, the following property is somewhat surprising. When technological opportunity (σ) is high, waves of activity will be *shorter*, to the extent that they exist at all. The reason is that when opportunity for invention is high, the industry will produce a steady stream of it, too steady to admit cycles. On the other hand (and less surprisingly), the *amplitude* of the deviations away from trend will be positively related to σ, since the randomness of the search process is positively related to it.

Consider now the probability of transiting between E and I. Let Q_T be the probability that the industry stays in I for an additional period, given that it has been there for T consecutive periods; the transition probabilities can be summarized by the matrix in Figure 2. While the first row is time-invariant, the second is not. Indeed, Equation (16) implies that $Q_T=0$ for $T>\sigma/2(\sigma-c\sqrt{2\pi})$, while for values of c close to $\sigma/\sqrt{2\pi}$ (which render extensive search a relatively unattractive option), it is easily shown that Q_1 is strictly positive. On average, therefore, Q_T is decreasing in T, and the escape probability from I therefore exhibits positive duration dependence.

Letting g_E and g_I denote the expected growth rates of the industry in its two states, we have:

$$(4.1) \qquad g_E = \sigma/\sqrt{2\pi} - c,$$
$$g_I = \sigma E[\max W(x_k') - w_k^*, 0].$$

Since the option of extensive search is always available, and since its return does not fluctuate, we certainly have $g_I \geqslant g_E$. Thus, one implication of our model is that the industry will grow faster during periods of intensive search. Since under Assumption 5 and the assumption of Theorem 2 neither state is absorbing, the long-run growth rate is a weighted average of g_I and g_E, the weights being the stationary-state probabilities.

Now compare two industries, A and B, both in the Schumpeterian region, which have the same c, but different σ's. Industry B will have higher growth (see (4.1)), and shorter implementation waves (Theorem 5).[10] This is a *long-run* implication, however, not to be confused with the short-run implication that waves of refinement (or imitation) represent periods of above-average growth.

4B. *Evidence on Specific Innovations*

Our theory emphasizes the supply side. It assumes that extensive search in some direction must precede further intensive exploration in that direction, and it analyzes the incentive aspects involved with the pursuit of these two kinds of activity. Kuznets (1962, esp. p. 22) found that the distinction between inventions on the one hand and improvements on those inventions on the other was a useful way to organize one's thinking about the growth of productive knowledge. But ultimately, if our dichotomy between discovery and its refinement is to be a useful one, one should be able to point to historical experience with particular inventions where this dichotomy makes sense.

One example is the artificial heart. The artificial heart program was started at the National Heart Institute with special congressional approval in 1963. This was indeed extensive search: A research program was aimed in a new direction; the main reason for it was the shortage of natural hearts. So far the only heart that the FDA has approved is the Jarvik heart, and although one recipient lived for 112 days following its implantation, everyone seems to agree that further improvement is needed and is likely.[11] The discovery of nylon also was the

[10] If we think of A and B as two firms, then there is some tentative support for this result in the work of Griliches (1986). Firm B faces greater technological opportunities than firm A and could thus be assumed to be spending more on "basic research." Griliches finds that productivity growth is indeed quite a bit higher in firms that do more basic research. Our view of causality here runs from higher σ on the one hand, to higher basic research and higher growth on the other. A referee pointed out, however, that this paper does not formalize basic research and more applied research, in the sense that both types of search in our model lead *immediately* to new ways of producing output in the subsequent period. This contrasts with the common perception that much of basic research is of no immediate commercial value. We therefore offer Griliches' evidence as merely suggestive.

[11] See the US Congress (1982).

outcome of a conscious decision by DuPont in 1928 to explore new dimensions, new chemical explorations.[12] Originally used for stockings, nylon has, through further refinements, been used in numerous other products since. The nuclear submarine too was, from the Navy's viewpoint, the outcome of search in a new dimension—alternative technologies such as the use of rechargeable batteries or carrying compressed air were unsatisfactory. In 1946 the US Navy made the decision to build the nuclear submarine which, because a nuclear reactor requires no oxygen, could stay underwater indefinitely.

Aside from instances in which a decision was made to provide a *particular* new direction, these are further examples of less *directed*, but still extensive search. These are instances in which elements are combined at random, such as arise in the pharmaceutical industry. The conventional method for creating a new drug starts with largely random sampling of components that show signs of having a biological impact such as slowing tumor cell production. Companies typically screen 5000 or more substances before finding a compound that is both safe and effective.[13] Recently, however, the advent of the computer (and computer graphics in particular) has led to a dramatic reduction in the cost of such search (in our model this amounts to a reduction in c). By simulating molecular interactions and generating images of molecules that might fit well with others, researchers are able to rule out a host of unpromising combinations without having to actually try them. Recently, computer graphics enabled researchers to predict that adding hydrogen to synthetic insulin molecules would lead to a smooth release of one insulin molecule at a time, and this will improve the treatment of diabetes. In our model, this reduction in c may push the pharmaceutical industry from Schumpeter's region into the region of i.i.d. growth (Figure 1), where it will grow faster and more evenly.

Three further sets of inventions can be argued to have had an element analogous to the extensive search that the model describes.[14] First, the search for superconducting materials has proceeded by trial and error, and it affords an excellent example of a successful extensive search (the somewhat accidental discovery at the IBM Zurich lab of ceramic oxide that superconducted at 90° K) followed up by further intensive search involving experimentation with other oxides by other investigators (who discovered that certain types of copper oxide were better still, by superconducting between 110 and 120° K).[15] Second, the attempt to determine where each human gene is located on the 23 pairs of chromosomes has so far proceeded essentially at random. Of the estimated 50,000–150,000 human genes, only 1200 or so have been located thus far. As in the pharmaceutical example discussed in the previous paragraph, however, recent advances in computers have dramatically reduced the cost of searching for genes and it is likely that the remaining genes will be discovered at a much

[12] Mueller (1962, p. 334). In fact, DuPont had a laboratory that was to be closed down; then it was decided to keep the lab open and allow it to be used for basic research.

[13] *NY Times* (1988, August 3).

[14] The paragraph elaborates on some helpful comments that a referee made.

[15] Hazen (1988).

faster rate.[16] Third, in his attempt to develop a working electric light bulb, Edison went through a process of essentially random testing of different materials for the filament. He started with carbon, but initially failed with it; he then tried platinum, chromium, silicon, tungsten, molybdenum, palladium, and boron, all without much success. Finally, when he acquired a better vacuum pump, he turned back to carbon (an intensive search) and it worked.[17]

Although our theory assumes a single consumption good, one might extend the interpretation to include product innovation—a firm decides whether to invent a new product (extensive search), or refine an existing one (intensive search). Gort and Klepper (1982) gathered data on dozens of product innovations, and compiled a list of innovations that took place during the lifetimes of these 23 products.[18] Our model would predict a tendency for the importance of the product improvements (as measured, say, by their effect on the growth-rate of the product's output) to successively decline—each successive intensive search is made from a distribution with a smaller variance. According to Gort and Klepper, innovations that occur during the early stages of the industry's development do indeed appear to be more important than the later ones.[19]

The instances given in this subsection are all examples of extensive search that, in several cases, was followed by intensive search. As such, these examples show that the distinction between extensive and intensive search is a useful one. But since these examples illustrate one-shot events, they do not constitute evidence of cycles or (except perhaps for the pharmaceutical example) perpetual extensive search. Documenting the latter would require one to look at the output (measured in terms of product and process innovations) of a particular type of research activity over an extended period of time.

5. EXTENSIONS AND CONCLUDING REMARKS

This paper derives a variant of Schumpeter's cycles from some minimal assumptions about technological knowledge. This is done in the simplest possible structure, with just two parameters: σ, which measures technological opportunity in both modes of search, and c, the cost of extensive search relative to intensive search. Technological opportunity, σ, here means something quite precise, which, given Assumptions 1–4, uniquely determines the entire distribution of payoffs for the economy in each mode of search. Such sharp conclusions follow from a bare-bones structure. What sort of further modifications could one look for? We end the paper with a series of remarks on possible extensions and other points of interest.

[16] US Congress, 1988.
[17] Friedel and Israel (1986).
[18] Gort and Klepper's Table 6 is of special relevance.
[19] Gort and Klepper (1982, p. 650).

5A. *Many Agents*

Ours is not a model of macro fluctuations. If one were to simply add more individuals to our economy, but keep their ideas isolated, the economy's growth rate would quickly converge to a constant, with no waves or cycles of any kind, and the fraction of resources devoted to extensive search would also converge to a constant.[20] Spillovers of knowledge are, however, pervasive in modern economies, and it is likely that the invention of something new would soon be followed by waves of applications and refinements. A careful analysis demands a model of diffusion of ideas among individuals or firms; a variety of approaches that could be taken is surveyed by Stoneman (1986). One would expect, however, that the slower the speed at which ideas spread from one agent to another, the longer it would take for the wave to work itself through. On the other hand, new basic inventions could arrive in the meantime and be added onto existing ones. Too much "mixing" of this sort would tend to eliminate or at least dampen the waves.

5B. *Variable Resources Devoted to Search*

Not only will the mix of resources devoted to applied and basic research vary over time, so will total research effort of the economy. Thus, at times when $g_I > g_E$, agents may still want to devote some resources to E, while when $g_I < g_E$, they might wish to devote even more. Members of a given generation would, in effect, choose the sample size. We have steered clear of this complication, because it can get quite involved (see Morgan (1983)). A serious treatment in the present context would need to introduce diminishing marginal utility of consumption so as to ensure an interior solution for the fraction of resources devoted to search.

5C. *Correlated* W_k

Following the Corollary to Lemma 1, we noted that the axioms said nothing about the possible correlation amongst the W_k. Independence maximizes ignorance—by knowing something about technology k, I learn nothing about technology j. But a theory of sustained periods of faster growth through learning may need to exploit such dependence. Suppose, for instance, that K is a subset of the integers, and that it is known that the W_k for $k \in K$ is a group of highly correlated technologies—if one works, then they all work. Then clearly, by finding a successful member of K, the economy has excellent growth prospects, as it can now turn to and sample all the other members of K. But how do agents come to believe that the W_k $(k \in K)$ are correlated? As a result

[20] One such model is in Jovanovic and Rob (1989). A completely different model of *macro* fluctuations associated with innovation, which is based on aggregate demand externalities is that of Shlaifer (1986).

of experience with prior technologies? Such questions could be pursued by allowing arbitrary, asymmetric correlation amongst the W_k in the prior, but, aside from the general observation that such economies will have higher serial correlations in growth rates,[21] not many other results are readily forthcoming.

5D. *Relation to the Multi-Armed Bandit*

New technologies are sampled infinitely often in this model. This is in contrast to the usual multi-armed bandit result (see Rothschild (1974) for a survey) that eventually the agent settles on one arm and pulls it forever. The reason for the difference between the results of the bandit formulation and our own is that the multi-arm bandit formulation bundles the consumption and investment decisions: to learn about arm k, one must consume the payoff it yields. As soon as one unbundles the two, new arms will be pulled infinitely often, and this is what the present formulation does.

5E. *Relation to Bayesian Analysis in General*

We follow the Bayesian approach to learning; the prior distribution on the functions $W_k(\cdot)$ is in our case just the Wiener measure discussed in Billingsley (1968). So long as Axioms (A.1)–(A.4) are imposed, it is thus possible to analyze optimal adaptive behavior even when prior information is minimal. In looking for axioms that support a unique prior, we have paid more attention than is customary to the process by which beliefs form.

Dept. of Economics, New York University, 269 Mercer St., New York, N.Y. 10003, U.S.A.

and

Dept. of Economics, University of Pennsylvania, Philadephia, PA 19104, U.S.A.

Manuscript received May, 1988; final manuscript received December, 1989.

REFERENCES

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. New York, NY: John Wiley and Sons.

CRESSIE, N. (1986): "Kriging Nonstationary Data," *Journal of the American Statistical Association*, 81, 625–634.

CRESSIE, N., AND R. HORTON (1987): "A Robust Resistant Spatial Analysis of Soil Water Infiltration," *Water Resources Research*, 23, 911–917.

FRIEDEL, R., AND P. ISRAEL (1986): *Edison's Electric Light: Biography of an Invention*. New Brunswick, NJ: Rutgers University Press.

GORT, M., AND S. KLEPPER (1982): "Time-Paths in the Diffusion of Product Innovations," *Economic Journal*, 92, 630–653.

GRILICHES, Z. (1970): "Productivity, R&D, and Basic Research at the Firm Level in the 1970's," *American Economic Review*, 76, 141–154.

HAZEN, R. (1988): "Perovskites," *Scientific American*, 258, 74–81.

[21] Because now the payoff to extensive search becomes random and autocorrelated.

Jovanovic, B., and R. Rob (1989): "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 56, 569–582.

Kuznets, S. (1940): "Schumpeter's Business Cycles," *American Economic Review*, 30, 257–271.

——— (1962): "Inventive Activity: Problems of Definition and Measurement," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, NBER. Princeton, NJ: Princeton University Press.

Mueller, W. (1962): "The Origins of the Basic Inventions Underlying DuPont's Major Product and Process Innovations, 1920–1950," in *The Rate and Direction of Inventive Activity*, NBER. Princeton, NJ: Princeton University Press.

Morgan, P. (1983): "Search and Optimal Sample Sizes," *Review of Economic Studies*, 50, 659–676.

Nelson, R. (1982): "The Role of Knowledge in R&D Efficiency," *Quarterly Journal of Economics*, 97 (1982), 453–470.

New York Times (1988): "Advances in Drugs, Courtesy of Computers," August 3, 1988.

——— (1988): "A Comeback for the Vacuum Tube," May 18, 1988.

——— (1988): "A Race to Make an Old Power Plant Idea Work," July 27, 1988.

——— (1988): "Scientists Revive a Lost Secret of Farming," November 22, 1988.

Rosenberg, N. (1972): *Technology and American Economic Growth*. White Plains, NY: Sharpe.

Rothschild, M. (1974): "Searching for the Lowest Price When the Distribution of Prices is Unknown," *Journal of Political Economy*, 82, 689–711.

Schumpeter, J. A. (1939): *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*, 2 vols. New York, NY: McGraw-Hill.

Shleifer, A. (1986): "Implementation Cycles," *Journal of Political Economy*, 94, 1163–1190.

Stoneman, P. (1986): "Technological Diffusion: The Viewpoint of Economic Theory," *Ricerche Economiche*, 40, 585–607.

Telser, L. G. (1982): "A Theory of Innovation and Its Effects," *Bell Journal of Economics*, 13, 69–92.

Wilson, R. (1975): "Informational Economies of Scale," *Bell Journal of Economics*, 6, 184–195.

U.S. Congress, Office of Technology Assessment (1982): "Case Study #9; The Artificial Heart, Risk and Benefit Analysis." Washington DC: Government Printing Office.

——— (1988): "Mapping Our Genes—the Genome Projects: How Big, How Fast?" OTA-BA-373. Washington DC: Government Printing Office.