LONG WAVES AND SHORT WAVES:

GROWTH THROUGH INTENSIVE AND EXTENSIVE SEARCH

by

Boyan Jovanovic
New York University

and

Rafael Rob
University of Pennsylvania

September 1987

ABSTRACT

This paper presents a model of growth through technical progress. The nature and scope of what is learned is derived from a set of axioms, and optimal search behavior by agents is then analyzed. Agents can search intensively or extensively. Intensive search explores a technology in greater depth, while extensive search yields new technologies. Agents alternate between these two modes of search. The economy grows forever and the growth rate is bounded away from zero. The growth rate is on average higher during periods of intensive search than during periods of extensive search. Epochs of higher growth are initiated by discoveries that call for further intensive exploration. This mechanism is reminiscent of the process described by Schumpeter as causing long-wave business cycles. Serial correlation properties of output and growth stem from the presence of intensive rather than extensive search. The two key parameters are technological opportunity, σ, and the cost of extensive search, c.

* We thank the C.V. Starr Center for Applied Economics for technical and financial assistance.

97

1. Introduction

This paper presents a model of growth through technical progress. The nature and scope of what is learned is derived from a set of axioms, and optimal search behavior by agents is then analyzed. Agents can search intensively or extensively. Intensive search explores a technology in greater depth, while extensive search yields new technologies. Agents alternate between these two modes of search. The economy grows forever and the growth rate is bounded away from zero. The growth rate is on average higher during periods of intensive search than during periods of extensive search. Serial correlation properties of output and growth stem from the presence of intensive rather than extensive search. The two key parameters are technological opportunity, σ, and the cost of extensive search, c.

Models in which growth does not eventually peter out have of late been of increased interest among economists. Recent examples are Lucas (1985), Romer (1987), and Prescott and Boyd (1987). The need for a more systematic look into the process of the growth of knowledge has motivated recent work by Nelson (1982), Telser (1982), Jovanovic and Rob (1987), as well as the work on learning when agents hold minimal prior beliefs about the structure, e.g. Frydman (1982), Marcet and Sargent (1986), and others.

The present paper tackles the same issues as the literature cited above, but it takes a different line of attack on the technological learning process that agents undergo. It is assumed that agents have a priori perfect knowledge of all available techniques that they might try out, but no knowledge of the level of output generated by them. Their past experience with some of these techniques enables them to make productivity inferences, and then make a decision as to which techniques would be suitable for future

26

experimentation. The state of knowledge is thereby constantly incremented. At the heart of our formulation is a set of probabilistic axioms regarding the "true" productivity of techniques, axioms that seem appropriate in nonparametric situations. What the agents see is not in violation of these axioms, and their behavior is optimal. Through their efforts to search for new technologies, long-run growth results.

The environment turns out to be a natural one for investigating the significance of a distinction long emphasized in the development literature (e.g., Rosenberg (1972)), the distinction between intensive and extensive search, as well as the distinction between applied and basic research. Intensive search (i.e., applied research) is the exploration in greater depth of a given technology, while extensive search (basic research) is the process by which new technologies are discovered. In our framework this distinction turns out to have empirical content: periods of intensive search result when a technology is discovered which it is optimal to explore further. The option to concentrate on extensive search is then passed over in favor of intensive search, because the latter yields higher growth. Thus the economy grows faster during such periods. This is roughly the intuition behind the result.

The point can be made in another way. In our structure, the expected returns to intensive search or applied research fluctuate over time, while the expected returns to extensive search or basic research are constant. One might justify this as follows: extensive search is the exploration of the unknown, and what is not known cannot be perceived to fluctuate. On the other hand, intensive search builds on the (random) outcome of extensive search, and conditional on the available information, its returns can be expected to vary over time. For instance, the discovery of nuclear power was the outcome of

99

basic research, or what we call extensive search. Once discovered, the rate of return to further development (i.e., intensive search) was raised above that to further basic research, and resources shifted to such more applied research. Thus our model predicts that periods in which agents concentrate on development research are high-return, high-growth periods, because agents are capitalizing on successes stemming from past efforts in basic research. When there are no such opportunities to be capitalized on, agents will focus on basic research, and the growth rate will be slower. This is intended as a model of an economy, or a firm, over time.

The implication that periods of intensive search or periods of heavier reliance on applied research are higher-growth periods can be tested using time-series data for countries or for firms. At the firm level, Scherer (1984, chapters 1 and 2) provides some anecdotal evidence in favor of this implication, but Griliches (1986) and Mansfield (1980) provide evidence against it. There is some question however, whether their measure of "basic R&D" also includes measures of what we have here termed "intensive search", and their results stem essentially from the cross-section variation over firms in basic R&D and growth-rates, and not from the time series fluctuations of these variables.

Major innovations such as the steam engine, railroads, electric power and nuclear power have been argued to be the cause of long waves in abnormally high economic activity, or long-wave business cycles (Schumpeter, 1939; Kuznets, 1940). The model captures this, and following the development of the model and its implications in section 2, we focus, in section 3, on some time-series properties of output, with special references to the long-wave hypothesis. Then, we derive results on the duration of epochs of intensive

100

search (an upper bound on their length is given), on the time-series properties of the growth rate, and on the matrix of transition probabilities between the state of extensive search and the state of intensive search. Search decisions turn out to depend only on the ratio σ/c; the first parameter measures technological opportunity and has a positive effect on the returns to both types of search, while the second is a measure of the costliness of extensive search relative to intensive search.

The fourth and concluding section briefly discusses the relation of this model to the multi-arm bandit model in particular, and to Bayesian analysis in general.

2. The model

There is a countable infinity of technology-types, and a continuum of each type. Let $x_i \in [0,1]$ be the level of use of technology of type i. There is a single output $z(x)$ associated with each technology-vector $x \in [0,1]^\infty$.

The function z is not known. At each date, t, there is a body of empirical knowledge, H^t (history):

$$H^t = (x^1, z^1, x^2, z^2, \ldots, x^t, z^t), \quad \text{where} \quad z^t = z(x^t).$$

Let $\Delta_n = \{x \in [0,1]^\infty | x = (x_1, \ldots, x_n, 0, \ldots.)\}, \quad \Delta = \bigcup_{n=1}^{\infty} \Delta_n, \quad x_1 \in \Delta_1.$

Search means choosing a vector x^{t+1}, and evaluating z at that vector. There are two types of search: intensive and extensive. Define $x|y_k = (x_1, \ldots x_{k-1}, y_k, x_{k+1}, \ldots x_n, \ldots)$.

101

Intensive Search: Let $n_t = N(H^t)$ be the smallest integer for which $\{x^1,\dots x^t\} \subseteq \Delta_{n_t}$. Choose a vector $x^\tau \in H^t$, a coordinate $1 \le k \le n_t$, and a value $x_k' \in [0,1]$. Interpretation: We are experimenting with a different level of use of the kth technology (which has been sampled before), to determine what impact it has on output, holding other components of x^τ constant. Then we observe $z(x^\tau|x_k') = z^{t+1}$. This experiment is costless.

Extensive Search

Pick a value $x_{n_t+1} \in [0,1]$, and a vector $x^\tau \in H^t$. Observe $z(x^\tau|x_{n_t+1}) = z^{t+1}$. Unlike intensive search, this is search in a new (i.e., (n_t+1)th) dimension. This type of search costs cZ_t, where $0 < c < 1$ and $Z_t = \text{Max}\{z^1,\dots z^t\}$. As we shall see, this amounts to assuming that each extensive search lead to a fraction c of current output being foregone.

Beliefs. The following axioms are imposed on the agents' beliefs. They will lead to a prior measure over the outputs of all technologies.

A1. Continuity. z is continuous in each variable separately. That is, each type of technology is given a locational context. As a consequence, optimal sampling within each dimension will be systematic, not random, as we shall see.

A2. 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.

A3. Constant proportional uncertainty. $\text{Var}\{z(x|x_k')|z(x) = z\} = \sigma^2(x_k'-x_k)z^2$ for $x_k' > x_k$. This makes the standard deviation of the output resulting from

102

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 at each new search. 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.

A4. <u>Independent increments</u>. 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 maximum ignorance. An increase, say, in output as one moves from x_k' to x_k contains <u>no</u> information on what will happen to output if we should experiment with x_k''.

<u>Remark</u>. We assume throughout that σ is known. 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.

<u>Lemma 1</u> (Billingsley (1968), p. 154): (A.1)-(A.4) imply that for each k, $[z(x)-z(x|0_k)]/z(x|0_k)$ is Brownian Motion with incremental variance σ^2, the percentage increase in output follows Brownian Motion in each technological dimension.

<u>Corollary</u>: The explicit representation of $z(\cdot)$ is

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

103

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

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

Remark: Eq. (1) says noting 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 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.

Choices available to agents.

There are overlapping generations that pass on information to each other. The transfer is free. Each generation can make just one search, and can then consume from the optimal technology hitherto sampled. Thus consumption and search investment are not bundled together. This is in contrast to the multi-armed bandit formulations in which it is assumed that the agent is forced to consume the payoff of the arm that he pulls. If z' is the output of the new technology that is sampled at t+1, then gross consumption (excluding the cost of search) is $\max(z', Z_t)$. Members of each generation are risk-neutral, and each generation consists of exactly one member. The only decision is whether to sample extensively or intensively, and, given the chosen mode of search, exactly which technology to sample. The payoff to each type of search will now be described in turn.

04

<u>Optimal extensive search</u>. If extensive search is the chosen option, then

<u>Theorem 1</u>. $x_{n_t+1} = 1$, and the expected payoff to extensive search is

$$Z_t(1 + \sigma/\sqrt{2\pi} - c) \tag{2}$$

<u>Proof</u>: 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

$$E\{Z_t(1 + \max(0, \sigma W_{n_t+1}) - c \,|\, x_{n_t+1}\} = Z_t[1 + \sigma x_{n_t+1}^{1/2}/\sqrt{2\pi} - c], \tag{3}$$

and 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

$$\sigma/\sqrt{2\pi} \geq c. \tag{B.1}$$

this assumption is maintained from this point on.

<u>Optimal intensive search</u>. 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 each interval of that partition are the values of z at its endpoints. This follows from (A.4).

Three stages are involved in intensive search. Stage 1: the agent selects a coordinate k, $1 \le k \le n_t$; 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.

Notationally, the setup is as follows. 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 the agent. Hence, $w_j^* \ge 0$. Prior to search at t, the agent can guarantee himself consumption

$$(4) \qquad Z = \prod_{j=1}^{\infty} (1+\sigma w_j^*).$$

Following an intensive search in dimension k at technology x_k' he gets

$$(5) \qquad \max(z',Z) = \prod_{j \ne k} (1+\sigma w_j^*)[1 + \sigma \max(W_k(x_k'),w_k^*)],$$

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

$$(6) \quad \pi(H^t) = \max_{1 \le k \le n_t} (\max_{x_k'} E[\max(z',Z)]) = \max_{1 \le k \le n_t} \prod_{j \ne k} (1+\sigma w_j^*)[1 + \sigma \max_{x_k'} E(\max[W_k(x_k'),w_k^*])].$$

Since Z_t is given at the start of the period, each generation will, in its decision about type of search, compare the expression in (2) (where Z_t is now given by (4)) with the expression in (6), and will choose extensive search if and only if

(6') $\quad (1+\sigma w_k^*)(1+\sigma/\sqrt{2\pi} - c) \;>\; 1 + \sigma \max_{x_k'} E\{\max[w_k^*, W_k(x_k')]\}, \quad 1 \le k \le n_t.$

(We have eliminated the multiplicative factor $\prod_{j\neq k}(1+\sigma w_j^*)$ which is common to (2) and (6).) We next investigate the RHS of (6)' and the optimal x_k'.

Stage 3: When learning along the kth dimension, we are learning about $W_k(\cdot)$, because of the multiplicative separability in eq. (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 optimal 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

$$(7) \qquad m = W^\alpha(\beta - x_k)/(\beta-\alpha) + W^\beta(x_k-\alpha)/(\beta-\alpha)$$

and with variance

$$(8) \qquad s^2 = (\beta - x_k)(x_k - \alpha).$$

In the sequel we shall need the following result:

107

Lemma 2: If $\epsilon \sim N(m,s^2)$ and $\bar{\epsilon}$ is a constant, then

$$u(m,s,\bar{\epsilon}) = E\max(\epsilon,\bar{\epsilon}) = m + (\bar{\epsilon}-m)F((\bar{\epsilon}-m)/s) + (s/\sqrt{2\pi})\exp(-(\bar{\epsilon}-m)^2/2s^2),$$

where F is the standard normal CDF.

Given the dimension k and an interval $[\alpha,\beta]$, it is evident that the maximization of (6) is equivalent to maximizing $u(m,s,w_k^*)$ subject to the constraints (7),(8). Let $v(\alpha,\beta,W^\alpha,W^\beta,w_k^*)$ be the maximized value of that program, i.e., the incremental percentage value of intensive search. Intensive search in $[\alpha,\beta]$ will take place if and only if

$$(10) \qquad 1 + \sigma v(\alpha,\beta,W^\alpha,W^\beta,w_k^*) > (1+\sigma w_k^*)(1+\sigma/\sqrt{2\pi} - c),$$

because by the discussion preceding (6)' the RHS is the incremental value of extensive search. Note that inequality (10) is time invariant, so that once a new dimension is explored, none of the previous dimensions will ever be further explored.

We now turn to characterizing those technologies which, upon their discovery, are developed further. These belong to the set

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

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

108

Theorem 2: (i) D is non-empty if and only if $c > \sigma/2\sqrt{2\pi}$.

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

$$(11) \qquad v[0,1,0,\omega,\max(0,\omega)] = (1+\sigma/\sqrt{2\pi} -c)\max(0,\omega) + 1/\sqrt{2\pi} - c/\sigma.$$

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' is then $1/2$, and eq. (8) yields $s^2 = 1/4$, lemma 2 yields $v(0,1,0,0,0) = \sigma/2\sqrt{2\pi}$, and the assertion follows. The "only if" part of (i) is demonstrated by looking at the derivatives of the 2 sides of (11). By applying the envelope theorem to v, we find that

$$(12) \qquad \partial v/\partial\omega = (1-F)x_k + I(\omega)F$$

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

$$(13) \qquad \partial\max(0,\omega)/\partial\omega = I(\omega).$$

Figure 1 makes it clear that if v is not above the RHS of (11) at $\omega = 0$, it can not exceed it for any ω, because for $\omega > 0$, the RHS of (12) is no greater than the RHS of (13) (which, in turn is equal to 1), and for $\omega < 0$, the RHS of (12) is non-negative while the RHS of (13) is zero. This proves the "only if" part of (i).

Turning to (ii), assume that D is non-empty. The existence of $\underline{w} < 0 < \bar{w}$ solving (11) will follow if we can show that (a) $\lim_{\omega\to-\infty} v = 0$ and (b) $\lim_{\omega\to\infty}[v - (1+\sigma/\sqrt{2\pi} -c)\omega] \leq 0$. Now (a) follows because $\lim_{m\to-\infty} u = 0$. For (b),

109

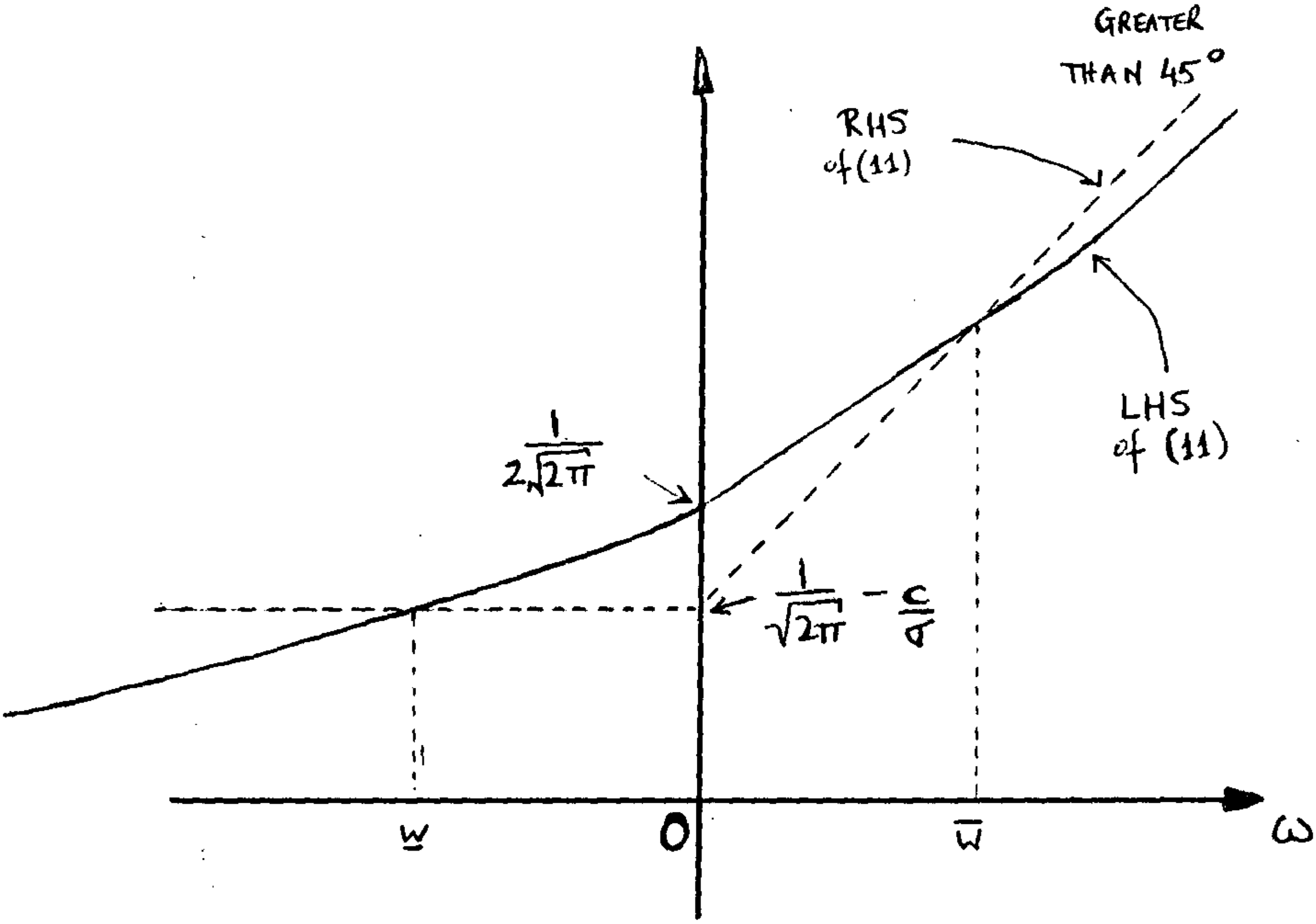
GREATER THAN 45°
RHS of (11)
LHS of (11)
$\frac{1}{2\sqrt{2\pi}}$
$\frac{1}{\sqrt{2\pi}} - \frac{c}{\sigma}$
$\underline{w}$
0
$\overline{w}$
ω

Figure 1

note that the derivative of v is no greater than 1 [see (12)], whereas the derivative of $(1+\sigma/\sqrt{2\pi} - c)\omega$ is strictly greater than 1 by (B.1). Q.E.D.

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

<u>Theorem 3</u>: In order for search to take place on an interval $[\alpha,\beta]$, we must have

$$(14) \qquad \beta-\alpha \geq 2(1 - c\sqrt{2\pi}/\sigma)(1+\sigma w^{*}),$$

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

<u>Proof</u>: The incremental value of intensive search on $[\alpha,\beta]$ is given by $v[\alpha,\beta,W^{\alpha},W^{\beta},\max(w^{*},W^{\alpha},W^{\beta})]$, which cannot exceed $v(\alpha,\beta,w^{*},w^{*},w^{*})$ (since v is increasing in W^{α} and W^{β} and since $w^{*} \geq \max(W^{\alpha},W^{\beta})$ by (9)). On the other hand, the incremental value of extensive search is $w^{*} + \sigma/\sqrt{2\pi}$. Furthermore, when $W^{\alpha} = W^{\beta} = w^{*}$, the optimal choice for x_k' is $(\alpha+\beta)/2$ which by (7), (8) and lemma 2 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

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

111

Hence, an intensive search on $[\alpha,\beta]$ is preferred to an extensive search <u>only if</u>

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

But this, by a slight rearrangement, is equivalent to (14). Q.E.D.

In particular, setting $w^* = 0$ (which by the assumption preceding equation (4) is smaller than the true w^*), we get a history-independent lower bound on the length of $[\alpha,\beta]$.

(15) $$\beta - \alpha \;\geq\; 2(1 - c\sqrt{2\pi}/\sigma).$$

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 (14) yields

(16) $$T \leq \sigma/2(\sigma - c\sqrt{2\pi}), \qquad \text{w.p.1.}$$

3. <u>The growth rate and long-wave business cycles</u>.

Although the true state space is H^t, it is helpful to think of the economy as being in one of two states: E = extensive search, and I = intensive search. Then if Q_T is defined to be the probability that the economy stays in I for an additional period, given that it has been there

112

for T consecutive periods, the transition probabilities can be summarized by the matrix

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

While the first row is time-invariant, the second is not. Indeed, eq. (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 economy in its two states, we have:

$$g_E = \sigma/\sqrt{2\pi} - c$$

$$g_I = \sigma E[\max W(x_k') - w_k^*, 0].$$

Since extensive search is always feasible, we certainly have $g_I \geq g_E$. Thus, one implication of our model is that the economy will grow faster during periods of intensive search. Since under (B.1) and the assumption of theorem 2

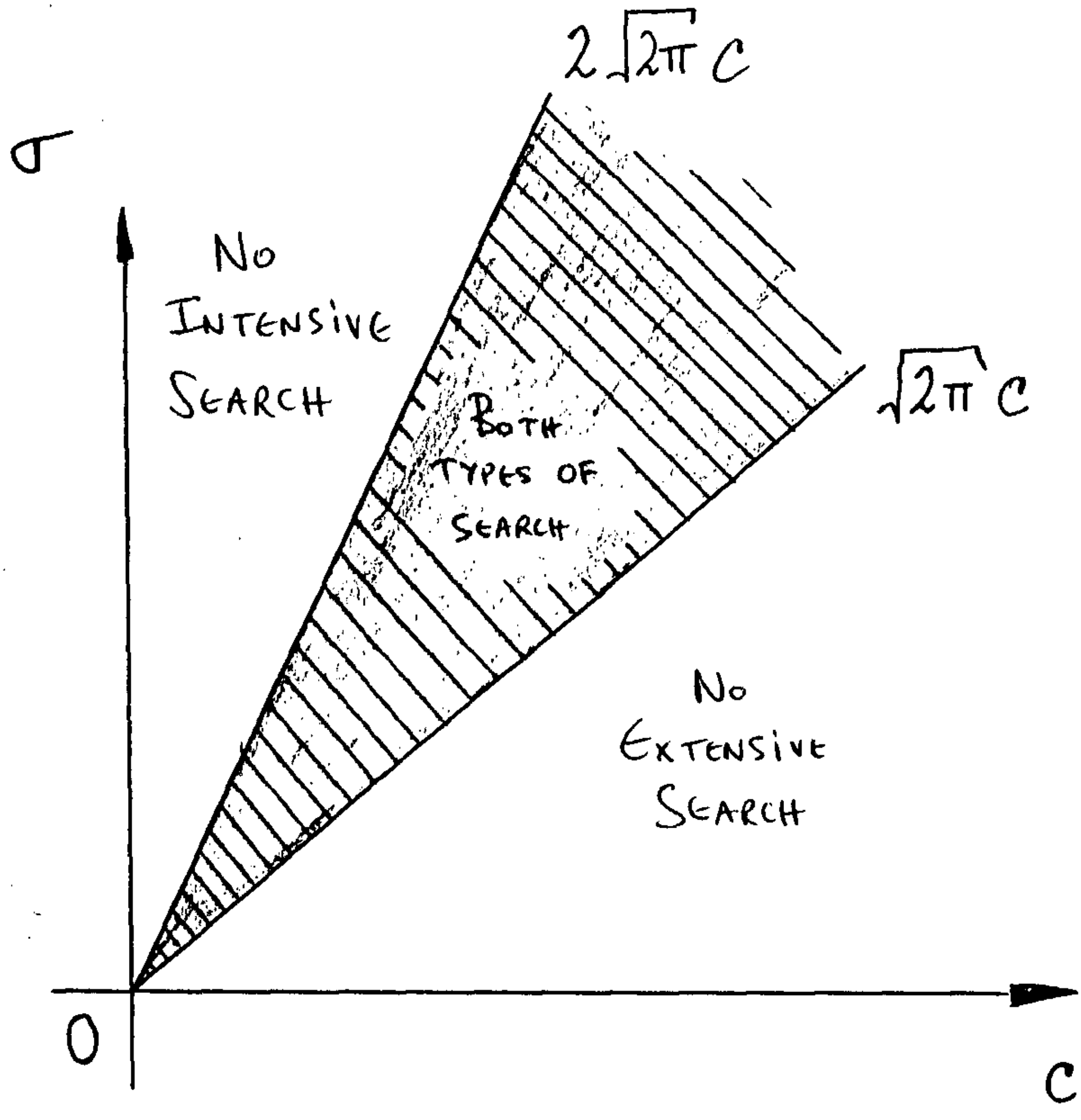
$2\sqrt{2\pi}\,c$
σ
No
Intensive
Search
Both
Types of
Search
$\sqrt{2\pi}\,c$
No
Extensive
Search
0
c

Figure 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.

Each uninterrupted spell in state I can be thought of as a "wave" of activity sparked by the discovery of a new technology. Not all technological discoveries lead to such waves: Only those technologies, k, with $W_k(1) \in D$ will 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)), and they can last at most $\sigma/2(\sigma - c\sqrt{2\pi})$ periods (see eq. 16). Figure 2 summarizes the parametric configuration necessary to produce long waves of activity. 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, whose shape is based on (B.1) and theorem 2(i).

If $c \to \infty$, the set D (defined just prior to theorem 2) becomes the entire line, so that the (1,1) cell of the above matrix becomes zero. The economy will never be in state E, and we are in the south-east region of figure 2. Moreover, if the economy 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 g_E. We would then be in the north-west region of the figure.

These predictions hinge on the assumption that the W_k are uncorrelated random functions. But, as we emphasized following Lemma 1, our axioms do not preclude the possibility of correlation amongst the W_k. Positive correlation between some of the W_k would significantly alter the model's implications. For instance, suppose that W_k and W_j are known to be positively correlated, and that agents experiment with technology j. If this turns out to be a failure, they will not try technology k but will turn elsewhere instead. But if j turns out to be a productive technology, they will then try k and will then expect an unusually high growth-rate of consumption. Thus the correlation of the W_k introduces (a) positive autocorrelation between the successive growth-rates during periods of extensive search, and (b) the possibility that during epochs of extensive search, growth on average will exceed that attained during epochs of intensive search. If, as Grilliches (1986) contends, (b) turns out to be the empirically-relevant case (in the time-series as well as in the cross-section), then further theoretical research ought to pursue the case where the W_k are positively correlated. If, however, one is to reconcile the positive serial correlation property of aggregate growth rates (Nelson and Plosser, 1982) with the apparent absence of serial correlation of growth rates at the firm level (Grilliches 1986, p. 152), one will, it appears, have to take into account rivalry amongst firms.

Returning to the Grilliches and Mansfield finding that faster-growing firms tend to do more basic research relative to applied research, we note that this is in accord with our model in the following basic sense. As noted in the previous section, without extensive search (basic research), the model yields a long-run growth rate equal to zero. Economies or firms that face (σ, c) pairs that allow faster growth will do more basic R&D. Whether they will also do more basic R&D relative to applied R&D will depend on how heterogeneous these firms or economies are with respect to the parameter c that they face.

116

4. Conclusion

We end by commenting on the relationship of our work to (a) multi-arm bandit analysis, and (b) Bayesian analysis in general. 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.

Although this paper makes minimal functional-form assumptions on the relationship between technologies and outputs, the approach we take has a well-defined probabilistic structure, based on the representation in (1). We have not abandoned the Bayesian approach to learning; the prior distribution on the functions $W_k(\cdot)$ is in each case just the Wiener measure discussed in Billingsley (1968). So long as certain axioms are imposed, it is thus possible to analyze optimal adaptive behavior even when prior information is minimal.

117

REFERENCES

Billingsley, P. Convergence of Probability Measures, New York: Wiley, 1968.

Frydman, R. "Towards an Understanding Market Processes, " American Economic Review, 72, no. 4 (September 1982): 652-68.

Grilliches, Z. "Productivity, R&D, and Basic Research at the Firm Level in the 1970's," American Economic Review, 76, no.1 (March 1986): 141-54.

Jovanovic, B. and R. Rob. "The Diffusion and Growth of Knowledge," unpublished, February 1987.

Kuznets, S. "Schumpeter's Business Cycles," American Economic Review, 30, no.2 (1940):257-71.

Lucas, R.E. "On the Mechanics of Economic Development," unpublished, May 1985.

Mansfield, E. "Basic Research and Productivity Increase in Manufacturing," American Economic Review, 70 (December 1980): 863-73.

Marcet, A. and T. Sargent. "Convergence of Least Squares Learning Mechanisms," October 1986.

Nelson, C. and C. Plosser. "Trends and Random Walks in Macroeconomic Times Series," Journal of Monetary Economics, 10, no. 2 (September 1982): 139-62.

Nelson, R. "The Role of Knowledge in R&D Efficiency," Quarterly Journal of Economics, XCVII, no. 3 (August 1982): 453-470.

Prescott, E.C. and Boyd, J. "Dynamic Coalitions: Engines of Growth," American Economic Review, (May 1987).

Romer, P. "Increasing Returns and Long-Run Growth," Journal of Political Economy, October 1986.

118

Rosenberg, N. Technology and American Economic Growth, Sharpe: White Plains, N.Y., 1972.

Rothschild, M. "Searching for the Lowest Price When the Distribution of Prices is Unknown," Journal of Political Economy, (June 1974).

Scherer, F. Innovation and Growth, MIT Press, Cambridge, Mass. 1984.

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

Shleifer, A. "Implementation Cycles," Journal of Political Economy, December 1986.

Telser, L.G. "A Theory of Innovation and Its Effects, " Bell Journal of Economics, 13, no. 1 (Spring 1982):69-92.

Wilson, R. "Informational Economies of Scale." Bell Journal of Economics 6 (Spring 1975): 184-95.

119