

ASYMMETRIC BUSINESS CYCLES: THEORY AND TIME-SERIES EVIDENCE

Daron Acemoglu

Department of Economics, Massachusetts Institute of Technology

and

Andrew Scott

Department of Economics, London Business School and All Souls College, Oxford

(Submitted to JME: August, 1995. This Version: June, 1997)

Abstract

We offer a theory of economic fluctuations based on intertemporal increasing returns: agents who have been active in the past face lower costs of action today. This specification explains the observed persistence in individual and aggregate output fluctuations even in the presence of i.i.d shocks, because individuals respond to the same shock differently depending on their recent past experience. The exact process for output, the sharpness of turning points and the degree of asymmetry are determined by the form of heterogeneity. Our general formulation, under certain assumptions, reduces to a number of popular state space (unobserved components) models. We find that on U.S. data our general formulation performs better than many of the existing econometric models, largely because it allows sharper downturns and more pronounced asymmetries than linear models, and is smoother than discrete regime shift models. Our estimates imply that only modest intertemporal returns are needed for our model to explain U.S. GNP, and that heterogeneity across agents plays an important role in the propagation of business cycle shocks.

Keywords: Asymmetries, Intertemporal Increasing Returns, Regime Shifts, Temporal Agglomeration, Unobserved Components.

JEL Classification: E32

This is a much revised version of our 1993 Working Paper, "A Theory of Economic Fluctuations: Increasing Returns and Temporal Agglomeration". We owe a special debt to James Hamilton whose insightful comments improved an earlier version of this paper. We also thank Philippe Aghion, Charlie Bean, Ricardo Caballero, Paul David, Steve Durlauf, Vittorio Grilli, Christopher Harris, Rebecca Henderson, Greg Hess, Per Krusell, Jim Malcomson, John Moore, Chris Pissarides, Danny Quah, Kevin Roberts, Donald Robertson, various seminar participants and an anonymous referee for useful comments. Remaining errors are the authors' responsibility.

1. Introduction

Aggregate economic fluctuations are characterized by successive periods of high growth followed by consecutive periods of low activity. The transitions between these periods of high and low growth are often marked by sharp turning points and considerable evidence suggests that at these moments the stochastic properties of the economy change and display asymmetries, see *inter alia*, Neftci (1984), Diebold and Rudebusch (1989), Hamilton (1989), Sichel (1993), Acemoglu and Scott (1994). The importance of tracking these movements in the business cycle is reflected in the considerable attention paid to a variety of coincident and leading indicators (e.g. Stock and Watson (1989), and the papers in Lahiri and Moore (1991)).

A natural way to model temporal agglomeration¹ and asymmetries in economic fluctuations is to assume non-convexities, such as discrete choice or fixed costs at the individual level, because such non-convexities imply that individuals concentrate their activity in a particular period. This implication has been analyzed with considerable success in the (S,s) literature. While the presence of fixed costs can account for the discreteness of economic turning points, it does not naturally lead to persistence because once an individual undertakes an action they are less likely to do so in the near future. Put differently, although the presence of fixed costs leads to increasing returns, these are *intra-temporal*; the full extent of economies of scale arising from fixed costs can be exploited within a period². As a consequence persistence in aggregate fluctuations relies on aggregation across heterogeneous agents: either more agents investing in the past increases the profitability of investment for others (e.g. Durlauf (1991) and (1993)) or aggregate shocks affect agents differently, leading to a smoothed response over time (e.g. Caballero and Engel (1991)).

This paper emphasizes an alternative explanation for persistent aggregate output fluctuations. In our model, there are *inter-temporal increasing returns* so that returns from an activity this period are higher if the activity occurred in the recent past. Therefore, an agent who was active in the recent past is more likely to be active now. We show how such a model explains a number of empirical features of business cycle fluctuations and also offers a framework which enables an economic interpretation of a number of unobserved component time series models of U.S. output.

Whether intertemporal increasing returns are important in propagating business cycles depends on evidence concerning two questions : are there important intertemporal linkages in firms' technology decisions? and do individual firms exhibit significant persistence in their activity and actions? The answers to these questions will vary depending upon the type of activity under consideration. In the case

¹ Following Hall (1991), we use this term to mean bunching of economic activity over time.

² To overcome this problem empirical implementations of (S,s) models sometimes include time-to-build considerations or decreasing returns at high levels of investment, e.g. Caballero and Engel (1994).

of radical changes to the capital stock, e.g. Cooper and Haltiwanger's (1993) automobile retooling, persistence is unlikely to be important. On the contrary, a radical change reduces the likelihood of another radical change tomorrow. However, Section 2 surveys micro evidence that firm level investment is highly persistent, and discusses the less formal but still informative findings from technology studies, the management science literature and organizational theory which support the notion that many important “qualitative” decisions (e.g. investment in new technology, product development, innovation, maintenance) exhibit intertemporal increasing returns. These findings motivate the model of Section 3 in which a firm has to decide each period whether to undertake both maintenance and investment. Maintenance has two effects: (i) increasing the productivity of existing technologies, and (ii) facilitating the adoption of new innovations. The interaction of these two roles leads to intertemporal increasing returns: firms find it profitable to maintain the newly adopted technologies and this in turn reduces the costs of adopting future innovations. As a result, investment costs are lower when the firm has invested last period, and a natural asymmetry is introduced in individual behavior: in response to a range of shocks, agents will find it profitable to invest only if they have invested in the recent past.

Sections 4 and 5 examine the aggregate economic fluctuations implied by individual level intertemporal increasing returns. We find that our model leads to a characterization of output dynamics where a cyclical component, which we may loosely think of as the “state of the business cycle”, plays a crucial role. This cyclical component can be highly persistent due to intertemporal increasing returns at the individual level, and can exhibit sharp turning points and significant *asymmetries*. Our model is tractable but also sufficiently general to capture rich nonlinearities. We show the cyclical component of output growth, which is crucial for business cycle fluctuations, follows a nonlinear autoregressive process as in the Smooth Transition Regression (STR) models of Granger and Terasvirta (1993). Different assumptions about the fundamentals of our economy lead to different forms of nonlinearities, thus to different STR specifications. This enables our model to nest a wide range of alternative time series models which have been used to model output growth. These econometric specifications differ in the extent to which they allow sharp turning points and asymmetries, and it is important to understand what underlies these differences.

Our model links the sharpness of turning points and the degree of persistence to the form of *heterogeneity*. For example, the symmetric “return-to-normality” model used in Harvey (1985), Watson (1986) and Clark (1987) is a special case of our model with idiosyncratic shocks drawn from a uniform distribution. Another extreme case of our model, when idiosyncratic shocks become very small relative to aggregate shocks, is the discrete regime shift model of Hamilton (1989) which exhibits very sharp turning points and pronounced asymmetries. When we turn to estimating our model in Section 6, we find that a good representation for U.S. data is somewhere in the middle of these two models. The

return-to-normality model does not capture the sharpness of downturns while the discrete regime shift model does not allow sufficiently for smooth transitions between different stages of the business cycle. Finally, both our own estimates and those of others in the literature give us another way of investigating the plausibility of our model. From these estimates, we calculate what the size of the required increasing returns are, and how large the variance of idiosyncratic shocks should be relative to aggregate shocks. We find that for our model to match U.S. business cycle fluctuations modest amounts of increasing returns and only a small variance of aggregate shocks relative to the variance of idiosyncratic shocks are sufficient.

2. Individual Persistence and Intertemporal Increasing Returns

Temporal agglomeration is naturally associated with fixed costs and qualitative choices, for example whether to perform a certain activity or make an indivisible investment. In the standard case, such activities are bunched within a period of time due to fixed costs, essentially because fixed costs imply the existence of *intra-temporal* increasing returns to scale. This observation lies at the heart of (S,s) models (e.g. Scarf (1959)) and implies that a brief period of activity is followed by periods of inactivity at the individual level. While (S,s) models receive support from the data (e.g. Bertola and Caballero (1990), Doms and Dunne (1994), Cooper, Haltiwanger and Power (1994)), there is also substantial evidence that firm level investment decisions are characterized by significant persistence. For example, using U.K. firm level data, Bond and Meghir (1994) find significant autoregressive effects in investment behavior. They estimate equations for the investment-capital ratio and find

$I_t/K_t = \alpha + 0.856 I_{t-1}/K_{t-1} + \beta Z_t + e_t$ where Z_t is a vector of firm relevant variables and e_t is a white noise disturbance. Evaluating the quadratic term at its sample mean yields an AR(1) coefficient of around 0.75. Bond et al (1994) estimate AR(1) and AR(2) models of investment using Belgian, French, German and U.K., firm panel data and find strongly significant investment lags, with the sum of the autoregressive coefficients around 0.3. Even the evidence in Doms and Dunne (1994), often used to support (S,s) models, reveals that the majority of firms have significant investments in most years of the sample, and that concentrated investment bursts are spread over several years. These findings suggest that there may exist intertemporal linkages as well as the fixed costs leading to intratemporal economies of scale.

The most obvious form of intertemporal economies of scale is learning-by-doing. More explicitly, consider the case where incorporating new knowledge is a slow and costly process, limiting the degree to which productive investments can be undertaken within a period. However, the more familiar an individual is with the most recent technology vintages, the cheaper it is to adopt the latest version. In contrast an individual who is not using the most recent vintage faces compatibility problems

when dealing with the frontier technology. The result will be that only limited innovative moves are taken within each period, with forward steps more likely to come from active agents. Empirical studies of technological innovation show that it is precisely this type of incremental changes that account for the majority of the productivity improvements (e.g. Abernathy (1980), Myers and Marquis (1969) and Tushman and Anderson (1986)).

More important for our paper, there is also a consensus that incremental innovations are more likely to come from firms who have been active in the earlier stages of product development. In Freeman's (1980, p.168) words *"the advance of scientific research is constantly throwing up new discoveries and opening up new technical possibilities, a firm which is able to monitor this advancing frontier by one means or another may be one of the first to realize a new possibility"*. Arrow (1974) and Nelson and Winter (1982) also emphasize the advantages possessed by incumbent innovators in being able to further cope with incremental changes. Abernathy (1980, p.70), using evidence from diverse industries, notes that *"Each of the major companies seems to have made more frequent contributions in a particular area"* suggesting that previous innovations in a field facilitate future innovations. One possible explanation of these findings are fixed effects: some firms may simply be good at innovating in certain areas. However, the industry wide work of Hirsch (1952), Lieberman (1984) and Bahk and Gort (1993) suggests that more than just individual fixed effects is operating. In the remainder of this paper we shall focus on this form of investment and its implications for aggregate fluctuations.

3. Individual Behavior in the Presence of Intertemporal Increasing Returns

(i) *The Environment*

We assume that firms (agents) are risk-neutral and forward looking and maximize profits. Each period a new technology becomes available which has a stochastic productivity that is revealed at the beginning of the period. The agent decides whether to adopt this technology or not. If the technology is adopted, the productivity of the agent increases permanently, starting from the current period³. To obtain the highest return from this innovation, its compatibility with existing technologies needs to be monitored. In particular, at the end of the installation period there is the option to gain additional productivity through maintenance⁴. Under these assumptions the firm's output is

$$y_t = y_{t-1} - \alpha_0 + (\alpha_1 + \alpha_2 s_t) - \alpha_2 s_t (1 - m_t) \quad (1)$$

where α_0 , α_1 and α_2 are positive parameters, and s_t and m_t are binary decision variables that equal 1 if

³ Time-to-build considerations can easily be incorporated in the return function and only serve to change the timing of returns.

⁴ See Pennings and Buitendam (1987) for the importance of maintenance type activities.

investment (s_t) and maintenance (m_t) are undertaken this period, and equal zero otherwise. If the new technology is adopted ($s_t=1$), the productivity of the firm is permanently higher. When there is no maintenance effort at the end of the period ($m_t=0$), this increase in productivity is not as large as it could be ($\alpha_1 + u_t$ instead of $\alpha_1 + u_t$). Deterministic depreciation is denoted by α_0 . We assume that u_t is a serially uncorrelated random shock to the productivity of investment with distribution function $F(\cdot)$. Concentrating on i.i.d disturbances enables us to clearly illustrate the additional persistence and dynamics generated via the intertemporal linkages of our model. Maintenance costs are assumed to be equal to a positive constant, c_0 (i.e. $C_t^m = c_0 m_t$). In this model, maintenance also has an additional role: it reduces the cost of future investments. In particular, investment costs at time t are given by:

$$C_t = (\gamma_1 - \gamma_2 m_{t-1}) s_t \quad (2)$$

where both γ_1 and γ_2 are positive. When equipment is maintained the firm's investment costs next period are lower by the amount γ_2 . In terms of our computing example, if all existing bugs in the system are removed thanks to maintenance, then new software can be installed and used much more effectively in the next period.

In each period, the firm decides whether to invest ($s_t=0$ or 1) and whether to maintain ($m_t=0$ or 1). Denoting the discount factor by β and the per period return by $r(\cdot)$, we have:

$$r(s_{t+j}, m_{t+j}, m_{t+j-1}, y_{t+j-1}, u_{t+j}) = y_{t+j-1} - \alpha_0 + (\alpha_1 + u_{t+j}) s_{t+j} - \alpha_2 s_{t+j} (1 - m_{t+j}) - \gamma_0 m_{t+j} - (\gamma_1 - \gamma_2 m_{t+j-1}) s_{t+j} \quad (3)$$

and the maximization problem of the firm at time t is:

$$\max_{\{s_{t+j}\}, \{m_{t+j}\}} E_t \left\{ \sum_{j=0}^{\infty} \beta^j r(s_{t+j}, m_{t+j}, m_{t+j-1}, y_{t+j-1}, u_{t+j}) \right\} \quad (4)$$

subject to (1) and taking y_{t-1} , u_t , m_{t-1} as given. In period $t+j$, the state variables are m_{t+j-1} , u_{t+j} , y_{t+j-1} and the choice variables are s_{t+j} and m_{t+j} .

Whether the firm invests in period t depends on u_t . In contrast, the return to maintenance only depends on whether or not investment occurs. If no current investment is undertaken, the only benefit of maintenance is the potential cost reduction in the following period. Instead, if there is current investment, future productivity also increases by α_2 . As a result there are three possibilities regarding maintenance: (i) always maintain (ii) never maintain, (iii) maintain only when there is investment. Because we wish to make maintenance a decision of the firm rather than a fixed characteristic, we

concentrate on (iii). To this end we assume⁵:

Assumption A:

$$\beta\gamma_2 < \gamma_0 < \frac{\alpha_2}{1-\beta} \quad (5)$$

The first inequality can be understood by noting that α_2 is the maximum benefit from maintenance in the absence of current investment: if there is investment next period costs are lower by α_2 , otherwise there is no benefit. Therefore, the first part of the inequality implies that maintenance is not worthwhile just to obtain future cost savings. On the other hand, with current investment, the minimum gain from investment is the present value of the productivity increase due to maintenance, $\alpha_2/(1-\beta)$. Consequently the second part of the inequality states that even without future cost savings, it is profitable to maintain if there is current investment. It follows that when (5) holds, we can limit our attention to the case where the firm maintains only when it invests, thus $m_t = s_t$, and the per period return simplifies to:

$$r(s_{t+j}^s s_{t+j-1}^s y_{t+j}^s u_{t+j}^s) = y_{t+j}^s - \alpha_0 + (\alpha_1 + u_{t+j}^s) s_{t+j}^s - (\delta_0 - \delta_1 s_{t+j-1}^s) s_{t+j}^s \quad (6)$$

where $\alpha_0 = \alpha_0 + \alpha_1$ and $\alpha_1 = \alpha_1$. Assumption A therefore enables us to write our problem in a way which focuses on the intertemporal increasing returns arising from the interactive term $\alpha_1 s_{t+j}^s s_{t+j-1}^s$. Even though we arrived at (6) using our “maintenance” model, clearly there are other microfoundations which would lead to a similar profit function. The important ingredient is intertemporal increasing returns: profits from an activity must be higher when the firm has been active in the recent past. In the rest of the paper, the exact microfoundations of (6) do not matter since we will be working directly with (6).

(ii) Optimal Decision Rules

Using (6) we can define the firm’s value function $V(\cdot)$ as

$$V(y_{t-1}^s, s_{t-1}^s, u_t^s) = \sup_{s_t^s} \{ r(s_t^s y_{t-1}^s, s_{t-1}^s, u_t^s) + \beta E_t V(y_{t-1}^s - \alpha_0 + (\alpha_1 + u_t^s) s_t^s, s_t^s, u_{t+1}^s) \} \quad (7)$$

Solving the agent's optimization problem gives (see the appendix):

⁵ Assumption A is stronger than we require but simpler to understand than the necessary condition,

$$\beta\gamma_2 \int_{\omega_0 - \omega_1}^{\infty} dF(u) < \gamma_0 < \frac{\alpha_2}{1-\beta} + \beta\gamma_2 \int_{\omega_0 - \omega_1}^{\infty} dF(u) \quad , \text{ where } T_0 \text{ and } T_1 \text{ are constants defined below.}$$

Proposition 1: The value function

$$\begin{cases} s_t=1 & \text{and } V(y_{t-1}, s_{t-1}, u_t) = \phi_0 + \phi_y y_{t-1} + \phi_s s_{t-1} + \phi_u u_t \\ & \text{if } u_t \geq \omega_0 - \omega_1 s_{t-1} \\ s_t=0 & \text{and } V(y_{t-1}, s_{t-1}, u_t) = \phi'_0 + \phi_y y_{t-1} \\ & \text{if } u_t < \omega_0 - \omega_1 s_{t-1} \end{cases} \quad (8)$$

is the unique function that satisfies (7).

To understand the dichotomous nature of the value function, consider the case where the agent does not invest ($s_t=0$). Then, the disturbance u_t is irrelevant to future profits and it does not matter whether the firm invested/maintained last period and the value function is linear in y_{t-1} . However, if the firm invests the value function depends linearly upon y_{t-1} , s_{t-1} and u_t . Thus, the optimal choice of s_t is conditional on whether the investment shock, u_t , is above a certain critical value, $\omega_0 - \omega_1 s_{t-1}$. This critical value depends on s_{t-1} due to the intertemporal non-separability in the cost function: it is less costly to adopt the new technology at time t if the firm invested in period $t-1$.

The critical value for investment is determined by comparing the return to investment (that is (8) evaluated at $s_t=1$) with the return to not investing (which is (8) evaluated at $s_t=0$). Using the expressions derived in the Appendix and (8), this condition can be expressed as follows: the firm should invest if and only if

$$\left[1 - \beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) \right]^{-1} \times \left[\frac{1}{1 - \beta} \alpha_1 - \delta_0 + \beta \delta_1 \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) + \frac{\beta}{1 - \beta} \int_{\omega_0 - \omega_1}^{\omega_0} u_{t+1} dF(u_{t+1}) \right] + \delta_1 s_{t-1} + \frac{1}{1 - \beta} u_t \geq 0 \quad (9)$$

This inequality therefore determines, the coefficients T_0 and T_1 in (8) (see (A3) in the Appendix). In particular, setting (9) equal to 0, with $s_{t-1}=0$ and then $s_{t-1}=1$ gives two equations which can be solved for T_0 and T_1 . The intuition behind this expression is a also good way of illustrating the main features of our model. The firm is comparing the strategy $s_t=1$ with $s_t=0$ ⁶. If $s_t=1$ production increases by $\alpha_1 + u_t$ for all periods compared to $s_t=0$, which has a net present value of

$$(1 - \beta)^{-1} (\alpha_1 + u_t) - (\delta_0 - \delta_1 s_{t-1}) \quad (10)$$

Any further benefits from choosing $s_t=1$ depend upon future values of u_t . There are three

⁶ Our results are unchanged if m_t and s_t lie in the interval $[0,1]$ rather than take discrete values. In this case, if (9) holds as a strict inequality agents choose one corner, $s_t=1$ and $m_t=1$; if (9) is strictly negative, $s_t=m_t=0$. If (9) holds as an equality agents are indifferent between any choice that has $m_t=s_t$. If in this case we impose that the agent chooses $s_t=m_t=1$ our results hold exactly.

possible cases:

(i) If $u_{t+1} \mathbf{O}(-4, T_0 - T_1)$, the agent will not invest regardless of s_t and there are no consequences beyond (10).

(ii) If $u_{t+1} \mathbf{O}(T_0 - T_1, T_0)$, the firm's $t+1$ investment decision depends upon s_t . The shock is only favorable enough for investment if the firm benefits from cost savings arising from past investment, in other words, $s_{t+1}=1$ only if $s_t=1$. In this case, investing *today* means a difference in expected discounted value next period of

$$\beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) \times \left\{ (1-\beta)^{-1} \left(\alpha_0 + \int_{\omega_0 - \omega_1}^{\omega_0} u_{t+1} dF(u_{t+1}) \right) - (\delta_0 - \delta_1) \right\} \quad (11)$$

where the first integral represents the probability that $u_{t+1} \mathbf{O}(T_0 - T_1, T_0)$ and the second is the expected value of u_{t+1} conditional on $u_{t+1} \mathbf{O}(T_0 - T_1, T_0)$. If both u_{t+1} and u_{t+2} fall in the region $[T_0 - T_1, T_0)$, the same additional benefit accrues in $t+2$. In other words, s_{t+2} only equals 1 when $s_{t+1}=1$, which in turn will only be the case when $s_t=1$. As $\{u_t\}$ is an i.i.d sequence, this additional benefit at $t+2$ is (11) multiplied by $\$ \text{Prob}(T_0 - T_1 < u_{t+1} \# T_0)$. A similar logic holds for all future periods, and summing these terms over time yields:

$$\left\{ 1 - \beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) \right\}^{-1} \times \left[\left\{ \beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) \right\} \times \left\{ \frac{\alpha_1}{1-\beta} + \delta_1 - \delta_0 \right\} + \frac{\beta}{1-\beta} \int_{\omega_0 - \omega_1}^{\omega_0} u_{t+1} dF(u_{t+1}) \right] \quad (12)$$

Equation (12) is the expected present value of future investments conditional on investing today. In particular, if it does not invest today, it will not invest in the future with shocks in the interval $[T_0 - T_1, T_0)$. This reasoning illustrates that there is an important difference in the way firms respond to investment shocks in high and low activity states. In other words, the marginal propensity to invest varies between these states. This state dependence relies entirely on $T_1 > 0$, which from (A3) in the appendix is equivalent to $\alpha_1 > 0$. Thus intertemporal increasing returns are responsible for this differential pattern of responses.

(iii) Finally, if $u_{t+1} \mathbf{O}(T_0, 4)$, agents invest regardless of whether they benefit from lower costs. However, while investment decisions are the same irrespective of s_{t-1} , costs are not. If $s_t = 1$ the cost of choosing $s_{t+1}=1$ is lower by the amount α_1 . This cost reduction has expected present value of $\$ \alpha_1 (1 - F(T_0))$. The same benefit accrues at $t+2$, if both $u_{t+1} \mathbf{O}(T_0 - T_1, T_0)$ and $u_{t+2} \mathbf{O}(T_0, 4)$, with expected value of $\$ 2 \alpha_1 [F(T_0) - F(T_0 - T_1)] [1 - F(T_0)]$, with similar expressions holding for $t+3$, etc. Summing over time gives:

$$\left\{ 1 - \beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1}) \right\}^{-1} \times \beta \delta_1 \int_{\omega_0}^{\infty} dF(u_{t+1}) \quad (13)$$

This expression represents the reduction in future costs arising from current investment and again reflects the persistence in s_t , captured by the integral between $\bar{T}_0 - \bar{T}_1$ and \bar{T}_0 .

The sum of (10), (12), and (13) is equal to (9) and characterizes the optimal decision rule of firms. The most important feature of this decision rule is the dependence of current actions on past decisions. Due to intertemporal increasing returns, shocks in the range $[\bar{T}_0, \bar{T}_0 - \bar{T}_1)$ lead to investment if received by an agent who has been active in the past ($s_{t-1} = 1$) but not for an agent who has not invested at $t-1$. This is the source of persistence in individual behavior.

4. Cyclical Fluctuations in the Aggregate Economy

(i) Characterizing Output Fluctuations

We now turn to the implications of individual level intertemporal increasing returns for aggregate economic fluctuations. We assume that the economy consists of a continuum of agents, normalized to 1, each facing the technology described above. We allow for heterogeneity across firms by assuming that firm i receives a shock $u_t^i = v_t + \epsilon_t^i$ where v_t is an aggregate shock and ϵ_t^i is a firm specific innovation. We assume that ϵ_t^i is drawn from a common distribution function $G(\cdot)$, with associated density $g(\cdot)$, and that v_t is i.i.d with distribution function $H(\cdot)$, and density $h(\cdot)$. Finally, we assume ϵ_t^i is uncorrelated across individuals and over time. Both shocks are normalized to have zero mean and are assumed to be observed before agents make their investment decisions.

The decision rule of the each firm is as in section 3: it will invest iff $u_t^i \geq \bar{T}_0 - \bar{T}_1 s_{t-1}^i$. Conditioning on the aggregate shock, it is optimal to invest if and only if:

$$\epsilon_t^i \geq \omega_0 - \omega_1 s_{t-1}^i - v_t \quad (14)$$

where \bar{T}_0 and \bar{T}_1 are derived from the distribution of u_t^i . Defining S_t as the proportion of agents that invest in period t (equivalently, the aggregate propensity to invest), we have:

Proposition 2:

Aggregate output follows the process

$$\begin{aligned}
\Delta Y_t &= \int \Delta y_t^i di = \alpha_0 + \alpha_1 S_t + \int_{i \in [s_t^i=1]} u_t^i di \\
&= \alpha_0 + (\alpha_1 + \nu) S_t + \int_{\omega_0 - \nu}^{\infty} \epsilon g(\epsilon) d\epsilon + S_{t-1} \times \int_{\omega_0 - \omega_1 - \nu}^{\omega_0 - \nu} \epsilon g(\epsilon) d\epsilon
\end{aligned} \tag{15}$$

where

$$\begin{aligned}
S_t &= \{1 - G(\omega_0 - \nu)\}(1 - S_{t-1}) + \{1 - G(\omega_0 - \omega_1 - \nu)\} S_{t-1} \\
&= \{1 - G(\omega_0 - \nu)\} + \{G(\omega_0 - \nu) - G(\omega_0 - \omega_1 - \nu)\} S_{t-1}
\end{aligned} \tag{16}$$

Equation (16) is crucial for the time series properties of aggregate output. S_{t-1} is the number (mass) of firms who invest at time $t-1$ and impacts directly on the number of firms that will invest at t , S_t . In particular, of the S_{t-1} firms who invested last period, those with an idiosyncratic shock greater than $\bar{T}_0 - \bar{T}_1 - \nu_t$ will invest now. This gives the number of firms investing in two successive periods as $(1 - G(\bar{T}_0 - \bar{T}_1 - \nu_t)) S_{t-1}$. In contrast, the $(1 - S_{t-1})$ firms who did not invest last period will be less willing to invest, and only those with an idiosyncratic shock greater than $\bar{T}_0 - \nu_t$ will do so, that is $(1 - G(\bar{T}_0 - \nu_t))(1 - S_{t-1})$ of them will be investing. Therefore, as shown in Fig. 1, S_t is a weighted average of points on the distribution function of idiosyncratic shocks $G(\cdot)$, where the location of these points depends upon the aggregate shock and the weights depend on S_{t-1} . Note that although (16) is crucial for the cyclical pattern of output, the dynamics of Y_t are more involved because output growth also reflects the non-zero average of the idiosyncratic shocks of all those agents currently investing (see (15)).

(ii) The Nature of Business Cycle Fluctuations

Proposition 2 outlines an unobserved components model for GNP. The law of motion for output growth consists of both a measurement equation, (15), and a state equation, (16). The state equation keeps track of the changes in the number of active agents which is an important determinant of aggregate output changes. Ideally, with data on both S_t and Y_t we would have a two equation system, but with S_t unobserved by the econometrician, we have an unobserved components model.

The number of agents investing in period t , S_t , can most naturally be interpreted as the cyclical component of output, or it can loosely be thought as the ‘‘state of the business cycle’’. In fact, variations in S_t not only alter the growth rate of output via (15), but also provide persistence, because S_t follows a time varying AR(1) process with autoregressive coefficient equal to $G(\bar{T}_0 - \nu_t) - G(\bar{T}_0 - \bar{T}_1 - \nu_t)$. Persistence is caused by shocks in the region $[\bar{T}_0 - \bar{T}_1, \bar{T}_0]$: in the case where $s_t^i = 0$, a value of u_{t+1}^i in this region implies that it is optimal for $s_{t+1}^i = 0$, whereas with $s_t^i = 1$, $s_{t+1}^i = 1$ would be optimal. Therefore, agents who

invested last period have a higher propensity to invest this period than those who did not. In the aggregate this implies that i.i.d shocks are converted into persistent cyclical fluctuations.

A distinctive feature of (15) and (16) is the time variation in the AR parameter in the state equation. In fact (16) implies that the cyclical component of output is described by a class of smooth transition regression (STR) model, see Granger and Terasvirta (1993). An STR model for an AR(1) process is of the form

$$x_t = (\eta_0 + \eta_1 M(x_{t-1})) x_{t-1} + \epsilon_t \quad (17)$$

where $M(x_t)$ is a continuous function. Different assumptions regarding $M(\cdot)$ imply different forms of STR models and so different degrees of smoothness and different types of nonlinearity. From this we can see that our model is an STR with $M(\cdot) = G(T_0 - v_t) - G(T_0 - T_1 - v_t)$. Therefore, the distribution of idiosyncratic shocks determine the form of the STR model and the nature of business cycle asymmetries. This finding, that the nature of business cycle dynamics depends on the form heterogeneity, will be a recurring theme in the rest of the paper. Thus one of the contributions of this paper is to establish a link between these popular state space models (both linear and k) and an economic model where the different stochastic properties have clear interpretations in terms of differences in economic fundamentals. To the best of our knowledge this paper represents the first such model. It can be observed that the cyclical component of output S_t , rather than observed output growth, is determined by a STR. This enables us to account for empirical findings of nonlinearities arising from business cycle asymmetries (as in the empirical evidence in Acemoglu and Scott (1994)), by linking the changing stochastic properties of output to the stages of the business cycle.

To see how (16) accounts for business cycle asymmetries note that the time varying AR coefficient means that the impact effect of aggregate shocks on output varies over the business cycle. Referring back to Fig 1, changes in v_t shift the position of the two points along the horizontal axis. An increase in v_t shifts the chord AB down, and S_t increases. However, the exact impact of v_t depends upon both the slope of the chord AB (which is determined by $G(\cdot)$ and v_t) and the weights on the two points (determined by S_t). As a result, the nonlinear autoregressive form of (16) is a source of path dependence in our model as well as persistence; a shock which changes S_{t-1} not only affects S_t through the AR coefficient but *also alters the way that the economy responds to future shocks* due to the interaction between v_t and S_{t-1} .

An interesting special case of our model is when idiosyncratic shocks, i.e. $G(\cdot)$, are uniform. It can be verified in this that (see section 6) the asymmetric interactions between S_t and v_t are absent, and the AR coefficient in (16) is constant. This implies that (15) and (16) approximate the standard "return-to-normality" state space model estimated on U.S. output by Harvey (1985), Watson (1986) and Clark (1987)⁷. More generally, when idiosyncratic shocks are non-uniform, (15) and (16) yield alternative

⁷ The uniform distribution case only approximates the return to normality model due to the non-normality of the measurement equation disturbance and because, if v_t is such that either $T_0 - v_t$ or $T_0 - T_1 - v_t$ is outside the support of ϵ_t , the autoregressive coefficient is no longer constant.

component based models which can account for a wide range of asymmetries and cyclical fluctuations. We will exploit this feature of our model in Section 6. For now however, we can note that although our model allows an important role for heterogeneity in determining business cycle properties, we do not need to monitor complicated changes in the cross-sectional distributions to keep track of the state of the economy: what matters for business cycles is not the exact position of each agent over time but the distribution function of idiosyncratic shocks around particular ranges. This makes the model tractable and easy to apply.

(iii) Determinants of the Time Series Properties of the Business Cycle

Even though the nonlinear nature of our model implies that there is no unique definition of persistence, a natural candidate is the degree of serial correlation in S_t conditional upon v_t :

$$\rho = \frac{\partial S_t}{\partial S_{t-1}} \Big|_{v_t} = G(\omega_0 - v_t) - G(\omega_0 - \omega_1 - v_t) \quad (18)$$

Recalling that \bar{T}_1 is an increasing function of the degree of intertemporal increasing returns, α_1^* , we obtain:

Corollary 1: An increase in α_1^* increases the degree of persistence, ρ .

Persistence is driven by fact that some agents have shocks in the interval $[\bar{T}_0 - \bar{T}_1, \bar{T}_0)$ and invest in this period only because investment costs are lower due to recent high activity. As \bar{T}_1 determines the measure of these marginal agents, and is itself an increasing function of α_1^* , serial correlation is strengthened when intertemporal increasing returns are higher.

To illustrate how the nature of the business cycle varies with different degrees of increasing returns consider the following simulation. Assume aggregate and idiosyncratic uncertainty to be equally important with both having a variance of 0.25, the former being normally and the latter uniformly distributed, and let the gains from learning-by-doing, α_1^* , be equal to 1.52 (see Section 6.1 for a justification of this choice). Figs. 2 and 3 show the cyclical component arising from these assumptions for the case $\alpha_1^*/\alpha_0^* = 1/3$ and $1/2$. Given the strong path dependence in our system, Figs. 2 and 3 are drawn for the same (suitably scaled) sequence of random shocks. Both figures illustrate how intertemporal increasing returns convert i.i.d shocks into cyclical fluctuations. Yet, the cyclical indicator is far more persistent in Fig.3 than in Fig. 2. The increased learning-by-doing persuades firms to

continue to invest even in the presence of mediocre productivity shocks, significantly reducing the noise in the cyclical component.

To understand the impact of the distribution of idiosyncratic shocks on the behavior of the cyclical component, we turn to a more general measure of persistence than (17), which was defined conditional on a given value of v_t . Integrating across all possible values of the aggregate shock, we arrive at a global measure of persistence;

$$P = \int \frac{\partial \mathcal{S}_t}{\partial \mathcal{S}_{t-1}} \Big|_{v_t} h(v_t) dv_t \quad (19)$$

Taking a first-order Taylor expansion of P around the mean of v_t which is zero, we obtain

$$P \approx G(\omega_0) - G(\omega_0 - \omega_1) + [g(\omega_0) - g(\omega_0 - \omega_1)] \int v_t h(v_t) dv_t \quad (20)$$

By definition the second term is zero, and if we take a further first-order Taylor expansion of $G(\cdot)$ around \bar{T}_0 , P can be approximated by $g(\bar{T}_0)\bar{T}_1$. Thus we can state:

Corollary 2: For given \bar{T}_0 and \bar{T}_1 , an increase in $g(\bar{T}_0)$ will increase P .

The intuition behind this corollary is once again related to the fact that individual persistence arises from shocks in the region $[\bar{T}_0, \bar{T}_0 - \bar{T}_1]$. In the aggregate economy, the distribution of idiosyncratic shocks is important because it determines the density of agents who are around this critical region. $g(\bar{T}_0)\bar{T}_1$ is a measure of the number of such marginal agents, so that the higher is $g(\bar{T}_0)$ the greater is serial correlation. Corollary 2 therefore implies that *spreads* of $g(\cdot)$ around \bar{T}_0 (which will often be produced by increases in heterogeneity, represented by increases in the variance of v_t) will reduce persistence to the extent that they lower the number of agents in the region $[\bar{T}_0, \bar{T}_0 - \bar{T}_1]$. Therefore cross sectional considerations exert an important influence on the stochastic nature of business cycles.

To analyze the impact of aggregate uncertainty on the business cycle we take a second-order Taylor expansion of (18) around 0 followed by an additional first-order Taylor expansion around \bar{T}_0 . This gives:

$$\begin{aligned} P &\approx G(\omega_0) - G(\omega_0 - \omega_1) + [g'(\omega_0) - g'(\omega_0 - \omega_1)] Var(v_t) \\ &\approx G(\omega_0) - G(\omega_0 - \omega_1) + g''(\omega_0)\omega_1 Var(v_t) \end{aligned} \quad (21)$$

Therefore, we have;

Corollary 3: Increases in the variance of aggregate shocks reduce (increase) the persistence of the cyclical component if $g(\cdot)$ is concave (convex) around T_0 .

Corollary 3 states the surprising result that for a large class of idiosyncratic distributions, the more volatile are aggregate investment shocks, the less important is the business cycle -- in the sense that the cyclical component becomes less persistent and aggregate output fluctuations are increasingly driven directly by v_t and not by the state equation. The intuition behind this is that when $g(\cdot)$ is concave around T_0 , increased volatility of the aggregate shock leads to the critical investment threshold being located at points with low density. In contrast, in the case where $g(\cdot)$ is convex around T_0 , more aggregate variability will take us to values of the density function that are on average higher and will increase serial correlation due to the increased weight of marginal agents.

(iv) Structural Heterogeneity

The purpose of this subsection is to show that when we extend our model to allow for different types of heterogeneity across agents, the importance of the cyclical component and the degree of nonlinearities associated with the business cycle may be enhanced. In particular we show that increasing the dispersion of agents may increase the amount of persistence provided by our cyclical indicator. Given that $\{S_t\}$ represents the extent of co-movement between agents, this is a surprising result which re-iterates the limitations of representative agent models.

We have so far only considered what Caballero and Engel (1991) call stochastic heterogeneity, that is heterogeneity in the form of idiosyncratic shocks. We now focus on *structural heterogeneity* by allowing investment cost functions to be firm specific. We assume agent i has investment costs given by⁸

$$C_t^i = (\delta_0^i - \delta_1 s_{t-1}^i) s_t^i \quad (22)$$

Using the solution outlined in Section 3, each agent invests iff

$$\epsilon_t^i \geq \omega_0^i - \omega_1 s_{t-1}^i - v_t \quad (23)$$

where the distribution function of T_0^i is $F^i(\cdot)$ with support set U and is determined from the distribution of the cost parameter, ω_0^i . Increases in the degree of structural heterogeneity are equivalent to mean-preserving spreads of $F^i(\cdot)$. The law of motion for S_t is now

⁸ The intertemporal increasing returns parameter, ω_1 , can similarly be made individual specific.

$$\begin{aligned}
S_t &= (1 - S_{t-1}) \left\{ \int_U (1 - G(\omega_0^i - \nu)) d\Gamma(\omega_0^i) \right\} + S_{t-1} \left\{ \int_U (1 - G(\omega_0^i - \omega_1 - \nu)) d\Gamma(\omega_0^i) \right\} \\
&= \left\{ \int_U (1 - G(\omega_0^i - \nu)) d\Gamma(\omega_0^i) \right\} + S_{t-1} \left\{ \int_U \{ G(\omega_0^i - \nu) - G(\omega_0^i - \omega_1 - \nu) \} d\Gamma(\omega_0^i) \right\}
\end{aligned} \tag{24}$$

Applying the same two successive Taylor expansions as in the last subsection, our global measure of persistence becomes

$$P \approx \int_U [G(\omega_0^i) - G(\omega_0^i - \omega_1)] d\Gamma(\omega_0^i) \approx \omega_1 \int_U g(\omega_0^i) d\Gamma(\omega_0^i) \tag{25}$$

Corollary 4: If $g(\cdot)$ is convex (concave), mean preserving spreads of $\Gamma(\cdot)$ increase (decrease) P .

Therefore, increased dispersion of agents in the form of greater structural heterogeneity can interact with stochastic heterogeneity to increase the persistence generated through the cyclical component. The intuition is that if $g(\cdot)$ is convex, the averages of neighboring points will be higher than $g(\cdot)$ itself. In particular, the higher is $g(\cdot)$ the higher is the measure of agents in the critical region where investment is only profitable when costs are low. Thus, the presence of structural heterogeneity leads to a more serially correlated process for $\{S_t\}$, and a mean-preserving spread of Γ , which is an increase in the degree of structural heterogeneity, increases persistence. Conversely when $g(\cdot)$ is globally concave, an increase in structural heterogeneity will reduce persistence. This result underscores the importance of heterogeneity in determining the form of economic fluctuations, and this theme will be pursued further in the next section.

5. Regime Shifts in Economic Fluctuations

One of the attractions of fixed cost models is that the discrete individual behavior they imply can explain the sharpness of business cycle turning points. While our general model implies that the cyclical component is serially correlated, it does not impose any conditions on the nature of turning points. A number of studies (e.g. Hamilton (1989), Acemoglu and Scott (1994), Suzanne Cooper (1994), Diebold and Rudebusch (1994)) have modeled output fluctuations by assuming the business cycle to be characterized by regime shifts, that is by abrupt moves from recession to expansion (or vice versa). In the next subsection we investigate the factors which determine the sharpness of turning points in our model. If moves from booms to recessions (and vice versa) are sharp, output dynamics can be well approximated by regime shifts models which have the advantage of being simple and quite parsimonious (e.g. Hamilton, 1989). In subsection (ii), we derive the discrete regime shift model as a special case of

our model when idiosyncratic heterogeneity disappears.

(i) *The Nature of Turning Points*

To analyze this issue we use our basic model with only stochastic heterogeneity, and focus on the distribution of S_t conditional on S_{t-1} , which is :

$$\rho(S|S') = \frac{h(v)}{(1-S')g(\omega_0 - v) + S'g(\omega_0 - \omega_1 - v)} \quad (26)$$

where $D(S^*S')$ denotes the density of S_t conditional on $S_{t-1}=S'$ and v_t is written as an implicit function of S_t via (16).

The extreme case of regime shifts corresponds to the case where $S_t=0$ or 1 so that $D(S^*S')$ has its mass concentrated at two particular points. More generally, if $D(S^*S')$ has marked peaks, then transitions between different states take on the character of regime shifts. Thus, if we can establish $D'(S^*S')$ is positive at $D(S^*S')=0$, we will have located a local minimum which implies $D(S^*S')$ cannot be single peaked. From (26) we have:

$$\text{sign}(\rho''(S|S')) = \text{sign}\left[\frac{((1-S')g''(\omega_0 - v) + S'g''(\omega_0 - \omega_1 - v))}{|(1-S')g'(\omega_0 - v) + S'g'(\omega_0 - \omega_1 - v)|} - \frac{h''(v)}{|h'(v)|}\right] \quad (27)$$

While no general statement can be made regarding the transition between states we have the following:

Proposition 3: The conditional density of S_t is non-unimodal if $g(\cdot)$ is more concave than $h(v)$ in the neighborhood of $D(S^*S')=0$.

Proposition 3 suggests that turning points tend to be abrupt when $g(\cdot)$ is locally more concave than $h(\cdot)$. Naturally when $g(\cdot)$ has its mass concentrated at a particular point, it will tend to be more concave. This can be seen in Fig. 1. If the idiosyncratic shock is uniformly distributed, $G(\cdot)$ is a straight line and output growth, though persistent, is distributed uniformly along the continuum (" 0 ", " 0 " + " 1 "). However, if the distribution function is concentrated in the middle (as in Fig. 1), economic states become more distinct in the sense that the conditional distribution of $\{S_t\}$ is concentrated in particular intervals of $(0,1)$. As a consequence, turning points are more likely to be well approximated by regime shifts.

To investigate this point further, we ran some simulations. Recall that Fig. 3 showed the case where $\sigma_1/\sigma_0=1/2$ and the aggregate and idiosyncratic shocks were equally uncertain with a variance of 0.25. Fig. 4 maintains the degree of increasing return at the same level but increases the variance of the

idiosyncratic shock to 2.5. The very different cyclical patterns in the two figures are readily apparent. In Fig. 3, turning points are extremely sharp, particularly the observations at around 37 and 73. In contrast, the greater importance of idiosyncratic shocks adds a considerable amount of noise to the cyclical indicator in Fig. 4. While the turning points at observations 37 and 73 can still be detected, they represent only two of several observations where the cyclical component changes direction.

(ii) *A Model of Regime Shifts*

Proposition 3 and related simulations suggest that the more concentrated the idiosyncratic distribution the more appropriate is a regime shift characterization. Therefore in this subsection we focus on a model where $u_t^i = v_t$ so that there is no idiosyncratic uncertainty. The interest of this special case is that it offers a theoretical justification for the widely used discrete Markov state space models (e.g. Hamilton, 1989).

Because the model only contains an aggregate shock, the laws of motion for the aggregate economy are the same as those for the individual firm. Let us introduce the notation

$$\begin{aligned} \text{Prob}[S_t=1 | S_{t-1}=1] &= p \\ \text{Prob}[S_t=0 | S_{t-1}=0] &= q \end{aligned} \tag{28}$$

Then from Section 3, we have:

Proposition 4: The stochastic process for the change in aggregate output is

$$\Delta Y_t = -\alpha_0 + (\alpha_1 + \nu)S_t \tag{29}$$

where S_t is a Markov chain with transition matrix

$$T = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \tag{30}$$

and

$$p = \int_{\omega_0 - \omega_1}^{\infty} dH(\nu) , \quad q = \int_{-\infty}^{\omega_0} dH(\nu) \tag{31}$$

As in the model with heterogeneity, cyclical fluctuations are the result of shocks being propagated by intertemporal increasing returns in a manner which requires a state space formulation for

output growth. The distinguishing feature of (29)-(31) is that fluctuations take the form of shifts between distinct economic states. In each of these states the economy behaves differently: not only does the growth rate differ ($\mu_1 \neq 0$) but if $p \neq 1 - q$ then so do the durations of booms and recessions. These are all common features with the popular model of Hamilton (1989) which is almost identical to (29)-(31).

This regime shift model also shares a number of similarities with Durlauf (1993) who explicitly models the transition between different states of the business cycle. In both models, there is a non-convexity and intertemporal increasing returns to scale with the end result being that white noise productivity shocks are converted into serially correlated output fluctuations. Both models also suggest a strong role for path dependence; in (29)-(31) shocks which shift the economy from one state to another are highly persistent as they affect the way the economy responds to future shocks. However, a key difference between ours and Durlauf's work comes in the form that the intertemporal non-separability takes. In Durlauf's model the intertemporal linkage arises through localized technological spillovers. In other words, firms have a higher propensity to invest if their neighbors invested in the recent past. Due to the externality Durlauf's model generates multiple long run equilibria in the sense that the stochastic process for output is non-ergodic. In contrast, in our model, because increasing returns to scale are internal, the process for $\{Y_t\}$ is ergodic and the equilibrium path is uniquely determined, i.e. given v_t and $\{S_{t-1}, S_{t-2}, \dots\}$, we know with certainty which state the economy will be in.

6. Econometric Evidence

In this section we first investigate whether large increasing returns are required at the micro level for our model to match the data by using the econometric results of other researchers. We then estimate a general form of our model to uncover the importance of asymmetries and the underlying heterogeneity in U.S. business cycles, and to compare the performance of our model to some existing econometric specifications.

(I) Regime Shift Models

Hamilton (1989) estimates a two state discrete Markov model for U.S. GNP similar to (29)-(31) and finds $p=0.9$, $q=0.76$ and $\mu_1=1.52$. To calculate the implied degree of intertemporal increasing returns we solve the equations in (A3) in the appendix using these estimated parameter values⁹. Fig. 5 shows different combinations of μ_1/μ_0 and F_v^2 that generate $p=0.9$ and $q=0.76$ when aggregate shocks are normally distributed. To obtain the same persistence in S_t , a higher variance requires more increasing

⁹ μ_0 does not influence the values of T_0 and T_1 and so (without loss of generality) it is set to ensure the model matches mean US output growth.

returns. This is because with a greater variance of shocks, agents are more likely to receive a future shock less than $\bar{T}_0 - \bar{T}_1$. This higher likelihood of a future switch lessens the expected future benefits from increasing returns and makes agents both less likely to invest and less likely to remain investing once they have started to do so. Thus with increasing returns of 22% ($\alpha_1/\alpha_0=0.22$), we obtain the appropriate values of p and q when $F_v^2=0.05$, and with $\alpha_1/\alpha_0=11\%$, we only need $F_v^2=0.01$. Therefore, a very small amount of intertemporal increasing returns is sufficient for our underlying economic model to match Hamilton's empirical findings. As well as providing persistence, intertemporal increasing returns generate considerable amplification of the productivity shock. For the case where $\alpha_1/\alpha_0=0.22$ and $F_v^2=0.05$, even though $F_v^2/\sigma_{\epsilon}^2 = 0.03$, the variance of Y_t is equal to 0.634. Relying *only* on a single productivity shock to drive output fluctuations, we need implausibly large learning-by-doing effects of 53% to explain Hamilton's results. But with an additional additive disturbance in (1), as assumed by Hamilton and all econometric implementation of unobserved component models, his results can be explained with very small amounts of intertemporal increasing returns and aggregate uncertainty (e.g. $F_v^2=0.01$ and $\alpha_1/\alpha_0=11\%$).

Results from alternative studies confirm this finding that only small scale intertemporal increasing returns are necessary to generate empirically observed regime shift behavior. Suzanne Cooper (1994) uses monthly industrial production from 1931 to estimate a transition matrix similar to (30). To match her estimates of the transition probabilities ($p=0.55$ and $q=0.46$) while also matching the variance of U.S. GNP growth (again without resorting to any additional productivity disturbances other than a unique aggregate shock), we need the saving in fixed costs to be only around 3% (i.e. $\alpha_1/\alpha_0=0.03$). Diebold and Rudebusch (1994) estimate equations analogous to (30) using U.S. industrial production (allowing for a time dependent \mathbf{T}). Assuming only one disturbance and using their estimated standard error for Y_t to calibrate F_v^2 (an underestimate as this implicitly sets $S_t=1$ for all t), we find that $\alpha_1/\alpha_0=0.8\%$ is sufficient to explain their results.

(ii) *The General Unobserved Components Model*

In this subsection we use (15) and (16) to obtain estimates of our structural parameters under alternative assumptions regarding the cross sectional distribution of idiosyncratic shocks. Estimating (15) and (16) with different idiosyncratic shock distributions enables us to compare our general model to some existing econometric specifications and also to perform a simple test of the importance of asymmetries in U.S. business cycles.

Assuming that idiosyncratic shocks are uniformly distributed over $[-a,a]$, (15) and (16) can be written as

$$\begin{aligned}\Delta Y_t &= b_0 + b_1 S_t + b_2 v_t + b_3 v_t^2 \\ S_t &= \left[\frac{a - \omega_0}{2a} \right] + \frac{\omega_1}{2a} S_{t-1} + \frac{v_t}{2a} \\ &= c + TS_{t-1} + u_t\end{aligned}\tag{32}$$

where the coefficients b_i are functions of the structural parameters T_0, T_1, a, ω_0 and ω_1 . Aside from the squared disturbance in the measurement equation, (32) is the standard return to normality model (e.g. Harvey (1989), ch. 3). Various versions of this model have been used to estimate the cyclical component of U.S. GNP, with Watson (1986) and Clark (1987) both estimating models similar to (32)¹⁰. If instead we assume that idiosyncratic shocks are distributed normally, we can take a second-order Taylor expansion and write^{11,12},

$$\begin{aligned}\Delta Y_t &= b_0 + b_1 S_t + b_2 v_t + b_3 v_t^2 + b_4 v_t S_{t-1} \\ S_t &= (1 - G(\omega_0)) + [G(\omega_0) - G(\omega_0 - \omega_1)] S_{t-1} + g(\omega_0) v_t + \\ &\quad (g(\omega_0 - \omega_1) - g(\omega_0)) v_t S_{t-1} \\ &= c + TS_{t-1} + K u_t S_{t-1} + u_t\end{aligned}\tag{33}$$

Equations (32) and (33) show that a simple test for asymmetry in either the state or measurement equation is to test the significance of the $v_t S_{t-1}$ term.

We estimate (32) and (33) using the growth rate of quarterly U.S. real GNP for the period 1954:1 and 1987:4. We augment the measurement equation in each case with a normally distributed additive measurement error. Estimation was performed using maximum likelihood via the Kalman filter. We also ignored the squared disturbance term which considerably simplified the estimation. In neither case did our estimates reveal any indication of heteroscedasticity, suggesting that dropping the squared term was not an important omission. Because in (32) and (33) the measurement and state

¹⁰ Watson and Clark estimate their model using the level of the U.S. real GNP allowing for a trend, a cycle and an irregular component. Thus our comparison is purely with their specification of the cyclical component, our equation (16). They both actually estimate the cyclical component as an AR(2) process. Acemoglu and Scott (1993) show this can easily be allowed for by letting intertemporal increasing returns operate with longer lags.

¹¹ Equation (32) is an approximation which is valid for any distributional assumption regarding idiosyncratic shocks. Different assumptions regarding $G(\cdot)$ lead to different estimates of the model's structural parameters.

¹² The exact expression for b_4 is as follows:

$$b_4 = [G(\omega_0) - G(\omega_0 - \omega_1)] \times \left[1 - \frac{\int_{\omega_0 - \omega_1}^{\omega_0} u g(u) du \times (g(\omega_0 - \omega_1) - g(\omega_0))}{[G(\omega_0) - G(\omega_0 - \omega_1)]^2} \right] + (\omega_0 - \omega_1) g(\omega_0 - \omega_1) - \omega_0 g(\omega_0).$$

This term disappears when idiosyncratic shocks are uniform.

equation have a correlated disturbance a simple alteration is required to the standard recursions of the Kalman filter (see Harvey (1989), p.112). Table 1 contains our estimation results, and Fig. 6 shows the (smoothed) estimates of S_t that emerge from the different assumptions on idiosyncratic disturbances. Both versions of the model suggest that a persistent cyclical component accounts successfully for serial correlation in U.S. output growth.

Examining the uniform distribution case (equation (32)), we find that the cyclical component is persistent, with an autoregressive coefficient of 0.52. The model also has a goodness of fit, $R^2=0.595$. The estimates of the structural parameters were $F_v^2 = 0.06$, and $F_s^2 = 1.45$ implying that the variance of idiosyncratic shocks is around 25 times that of aggregate shocks. This estimate is considerably larger than the ratio of idiosyncratic to aggregate shock variances that we will obtain from the estimation of (33), and the micro data estimates of Schankerman (1991) and Davis and Haltiwanger (1992) lie in between these two estimates. Given our results in the last section regarding the links between heterogeneity and regime shifts, this finding, both with uniform and normal distributions, implies that U.S. business cycles are considerably smoother than discrete regime shifts (recall that a regime shift model is a special case of both (32) and (33) with the variance of idiosyncratic shocks equal to zero). This finding therefore also suggests why statistical tests of Hamilton's (1989) model (e.g. Hansen (1992) and Garcia (1992)) reject the notion of regime shifts between distinct states in favor of smoother alternatives. We also recovered α_1 to be 0.0201, which implies that when the business cycle indicator S_t is at its peak, GNP growth is 2% higher compared to a trough. Finally, using our estimated structural parameters and assuming a real interest rate of 4% per annum we can use equations (A3) to calculate estimates for α_0^* and α_1^* , the learning-by-doing parameters. We find that $\alpha_0^* = 1.96$ and $\alpha_1^* = 0.54$ implying intertemporal increasing returns (α_1^*/α_0^*) of around 27.6% (as a proportion of fixed costs).

The estimates of equation (33) which allows for asymmetries reveal considerably greater persistence in the cyclical component -- the autoregressive coefficient in the state equation is now 0.67 as opposed to 0.52 in the uniform case. More importantly, the R^2 in this case goes up to 0.712, thus allowing for asymmetries enables us to explain an additional 12% of U.S. output growth fluctuations. As noted above, equation (32) is essentially what Clark, Harvey and Watson have estimated. It is therefore important that our general model that allows for asymmetries outperforms this "linear" model and that the interaction term between v_t and S_{t-1} is significant in both the state and the measurement equations. Fig. 6 shows the estimated sequence for the cyclical indicator arising from (32) and (33). The strong correspondence between the two sequences is encouraging because it implies that we are uncovering the same underlying component in both exercises. However, there is a major difference between the two series: the asymmetric model, (33), enables

much sharper downward swings into recession, and it is this feature of the data which accounts for the success of the more general model. These sharper downswings are captured by the positive coefficient on the $v_t S_{t-1}$ term and are in line with the findings of Neftci (1984) who used a different statistical methodology than the one here.

Uncovering the underlying economic parameters from these econometric results, we find $\alpha_1 = 0.017$ which is less than the estimate of the growth difference between booms and recessions that we obtained from (32). However, note that this does not imply growth to be higher by only 1.7% at the peak of the cycle because in contrast to (32), we also have the added flexibility of having the asymmetry term $v_t S_{t-1}$ contributing to growth. The presence of this term implies that as long as we remain in a boom (i.e. positive values of v_t), we obtain an added growth effect and this is of the order of 0.2%. Hence, at the peak of the cycle growth is higher by 1.9%. But also this term implies that when there is a downturn ($v_t < 0$), this can happen rather sharply, and is therefore the source of the superior performance of our general model over the linear returns-to-normality specification. Nevertheless, as noted above, even though downturns are sharp, they are considerably smoother than those implied by a discrete regime shift model.

Returning to the rest of the estimates, F_v^2 is 0.035 and F_ϵ^2 is 0.1207. So our estimate of the ratio of the variance of idiosyncratic shocks to aggregate shocks is around 3.5 gives an important role to idiosyncratic variability in shaping cyclical fluctuations, but is considerably less than the approximate estimate of 10 from Schankerman (1991) and Davis and Haltiwanger (1992), and that implied by the estimation of (32) above. Resorting again to (A3) we find that our asymmetric model leads to estimates of $\alpha_0 = 1.81$ and $\alpha_1 = 0.63$. Thus, intertemporal effects of around 1/3 are required to explain U.S. fluctuations. Since there are more parameters estimated in equation (33), our economic (structural) parameters are overidentified. Performing a Wald test on these restrictions gives a test statistic of 4.1 which is asymptotically distributed chi-squared with two degrees of freedom, thus comfortably accepting the overidentifying restrictions. This result suggests that our econometric specification and economic model provide a quite good representation for the cyclical dynamics of U.S. GNP.

7. Conclusion

We have outlined a theory of economic fluctuations based on internal intertemporal increasing returns in a model of discrete investment choice. This model is motivated by microeconomic findings of persistence in firm level investments as well as the emphasis placed in technology studies on the importance of learning-by-doing in the adoption of new production techniques. Incorporating these effects into a model of a firm's investment choice leads to a tractable

model of business cycle fluctuations.

This tractability enables us to fully analyze the determinants of aggregate economic fluctuations. Our theoretical findings are: (i) Intertemporal increasing returns naturally lead to temporally agglomerated and asymmetric economic fluctuations. This implies persistent periods of low and high growth separated by business cycle turning points even when the underlying shocks are i.i.d. (ii) Heterogeneity plays a key role in determining the extent of nonlinearities, asymmetries and the sharpness of turning points. (iii) Although heterogeneity plays a key role in the nature of cyclical fluctuations, all the business cycle relevant information is captured in one variable, the average number of active agents. Hence, our model enables a synthesis between representative agent models and those stressing the importance of heterogeneity.

Under certain simplifying assumptions, our model reduces to the popular unobserved component models of Watson (1986), Clark (1987) and Hamilton (1989) all of which place a special emphasis on an underlying cyclical indicator. An additional attraction is the model's ability to offer a general formulation for cyclical components, and provide a simple test for the presence of business cycle asymmetries. Our general model with asymmetries provides a good fit to U.S. business cycles. In particular it matches the pronounced asymmetries and the sharp downturns that cannot be captured by a linear model. We also find that even though there are sharp turning points, business cycles are considerably smoother than those implied by a discrete regime shift model. In the context of our model, this means that heterogeneity and idiosyncratic shocks play an important role in the propagation of business cycle shocks. Finally, estimates of the degree of internal intertemporal increasing returns necessary to account for U.S. business cycles are fairly modest. Adding other sources of uncertainty to the model or allowing for spillovers between agents would serve to reduce even further the extent of internal increasing returns required. Assessing the relative importance of internal and external intertemporal increasing returns is an obvious topic of further research. At this stage, our results lead us to conclude that intertemporal increasing returns may be an important channel of persistence, amplification and asymmetries in economic fluctuations.

Table 1 : Estimates of General and Uniform Model

	Uniform Case		Normal Case	
Parameter	Estimate	Tstatistic	Estimate	Tstatistic
b_0	0.0004	1.691	0.0020	1.957
b_1	0.0201	3.107	0.0173	3.284
b_4			0.1027	1.985
c	0.1826	2.232	0.1769	2.173
T	0.5213	5.669	0.6653	6.339
K			0.3229	3.496
F_u^2	0.0070	2.703	0.0065	2.842
R^2	0.595		0.712	

Sample period 1957:1,1987:4. Data = Quarterly growth in U.S GNP. Estimation by Maximum likelihood via Kalman filter.

Appendix

Proof of Proposition 1

We need to establish that (i) the value function defined in (8) satisfies (7) for particular values of \bar{T}_0 and \bar{T}_1 and \mathbf{B} 's (ii) the value function satisfies a transversality condition.

To establish (i) we substitute (8) into (7) to derive:

$$\begin{aligned}
 & \text{If } u_t \geq \omega_0 - \omega_1 s_{t-1}, \text{ then} \\
 & \phi_\alpha + \phi_y y_{t-1} + \phi_s s_{t-1} + \phi_u u_t = y_{t-1} + \alpha_0 + \alpha_1 + u_t - (\delta_0 - \delta_1 s_{t-1}) \\
 & + \beta \left[\int_{\omega_0 - \omega_1}^{\infty} (\phi_0 + \phi_y (y_{t-1} + \alpha_0 + \alpha_1 + u)) + \phi_s + \phi_u u_{t+1} \right] dF(u_{t+1}) \\
 & + \int_{-\infty}^{\omega_0 - \omega_1} (\phi'_0 + \phi_y (y_{t-1} + \alpha_0 + \alpha_1 + u)) dF(u_{t+1})
 \end{aligned} \tag{A1}$$

where a similar equation applies for $v_t < \bar{T}_0 - \bar{T}_1 s_{t-1}$. Equating coefficients gives

$$\begin{aligned}
 \phi_\alpha = \phi_u &= \frac{1}{1-\beta} & \phi_s &= \delta_1 \\
 \phi_0 - \phi'_0 &= \frac{\frac{1}{1-\beta} \alpha_1 - \delta_0 + \beta \delta_1 \int_{\omega_0 - \omega_1}^{\infty} dF(u_{t+1}) + \frac{\beta}{1-\beta} \int_{\omega_0 - \omega_1}^{\omega_0} u_{t+1} dF(u_{t+1})}{1-\beta \int_{\omega_0 - \omega_1}^{\omega_0} dF(u_{t+1})}
 \end{aligned} \tag{A2}$$

To find expressions for \bar{T}_0 and \bar{T}_1 we use (8) to see under what conditions the left hand side of (5) is greater evaluated at $s_t=1$ than evaluated at $s_t=0$. This gives

$$\begin{aligned}
 \omega_1 &= \delta_1 (1-\beta) \\
 \omega_0 &= [1-\beta \int_{\omega_0 - \delta_1 (1-\beta)}^{\omega_0} dF(u_{t+1})]^{-1} \times [(1-\beta) \delta_0 - \alpha_1 - \\
 & (1-\beta) \beta \delta_1 \int_{\omega_0 - \delta_1 (1-\beta)}^{\infty} dF(u_{t+1}) - \beta \int_{\omega_0 - \delta_1 (1-\beta)}^{\omega_0} u_{t+1} dF(u_{t+1})]
 \end{aligned} \tag{A3}$$

Defining the right hand side of the definition of \bar{T}_0 as $\succ(\bar{T}_0)$ we need to prove $\succ(\bar{T}_0)$ has a fixed point.

Defining $z(u)$ as $\succ(u)-u$ we have

$$\lim_{\omega \rightarrow +\infty} z(\omega) = -\infty \quad \text{and} \quad \lim_{\omega \rightarrow -\infty} z(\omega) = +\infty \quad (\text{A4})$$

Continuity of $z(u)$ follows from continuity of $\succ(\cdot)$ and by the intermediate value theorem $z(u)$ must have a zero. Thus a fixed point, \bar{T}_0 , of $\succ(\cdot)$ exists.

(ii) The value function defined in (8) satisfies a transversality condition iff

$$\lim_{t \rightarrow \infty} \beta^t \int_{-\infty}^{\infty} V(\alpha_{t-1}, s_{t-1}, u_t) dF(u_t) = 0 \quad (\text{A5})$$

since $V(\cdot)$ is linear condition is satisfied for $\beta < 1$.

Establishing uniqueness of \bar{T}_0 requires proving $z(\bar{T}_0)$ has a unique fixed point. $z(\bar{T}_0)$ is everywhere differentiable and its derivative equals

$$\begin{aligned} & \omega_0 f(\omega_0 - (1 - \beta)\delta_1) \times [1 - \\ & \beta)\delta_1 \int_{\omega_0 - (1 - \beta)\delta_1}^{\infty} dF(u_{t+1}) - \\ & -(1 - \beta)\delta_1][1 - \beta \int_{\omega_0 - (1 - \beta)}^{\omega_0} \end{aligned} \quad (\text{A6})$$

Substituting $z(\bar{T}_0) = 0$, we have $z'(\bar{T}_0) = -1$ which establishes that $z(\bar{T}_0) = 0$ can only be true at a unique value of \bar{T}_0 .

We now establish by contradiction that there can be no nonlinear solutions to the recursion (7). Let $W(y_{t-1}, s_{t-1}, u_t)$ be a solution to (7) and for given values of the state variables let s_t^w be the optimal choice of s_t . Thus

$$\begin{aligned} W(y_{t-1}, s_{t-1}, u_t) &= y_{t-1} + \alpha_0 + (\alpha_1 + u_t - \delta_0 - \delta_1 s_{t-1}) s_t^w \\ &+ \int_{-\infty}^{\infty} W(y_{t-1} + \alpha_0 + (\alpha_1 + u_t) s_t^w, s_t^w, u_{t+1}) dF(u_{t+1}) \end{aligned} \quad (\text{A7})$$

Observation 1: $s_t^w(y_{t-1}, s_{t-1}, u_t)$ cannot depend on y_{t-1} as returns from a higher y_{t-1} accrue under both $s_t^w = 0$ and $s_t^w = 1$.

We know by assumption that both $V(.,.,.)$ in (8) and $W(.)$ satisfy (7). Take a value of u_t where $V(.)$ gives $s_t^v=0$. The difference between (5) evaluated at $W(.)$ and $V(.)$, for given u_t , is

$$\begin{aligned} & W(y_{t-1}, s_{t-1}, u_t) - \Phi_0 - \Phi_y y_{t-1} = (\alpha_1 + u_t - \delta_0 - \delta_1 s_{t-1}) s_t^w + \\ & \beta \{ E_t W(y_{t-1} + \alpha_0 + (\alpha_1 + u_t) s_t^w, s_t^w, u_{t+1}) \\ & - \int_{\omega_0}^{\infty} (\Phi_0 + \Phi_y (y_{t-1} + \alpha_0) + \Phi_u u_{t+1}) dF(u_{t+1}) - \int_{-\infty}^{\omega_0} (\Phi_0' + \Phi_y (y_{t-1} + \alpha_0)) dF(u_{t+1}) \} \end{aligned} \quad (A8)$$

From Observation 1 this holds for all y_{t-1} for fixed u_t , implying that

$$W(y_{t-1}, s_{t-1}, u_t) - \beta E_t W(y_{t-1} + \alpha_0 + (\alpha_1 + u_t) s_t^w, s_t^w, u_{t+1}) = K_1 + k_1 y_{t-1} \quad (A9)$$

Thus $W(.,.,.)$ is linear in y_{t-1} with a constant coefficient. To see this note that we can repeat this exercise for $s_{t-1}=0$ or 1 and get the same coefficient k_1 . Similarly we can repeat the exercise for values of u_t at which $s_t^v=1$ and get the same relationship.

Observation 2: If

$$E_t \sum_{j=0}^{\infty} \beta^j r(y_{t+j-1}, s_{t+j-1}, u_{t+j} | s_t = 1) > E_t \sum_{j=0}^{\infty} \beta^j r(y_{t+j-1}, s_{t+j-1}, u_{t+j} | s_t = 0) \quad (A10)$$

is true for $u_t = u$, it is also true for $u_t = u + \epsilon$, such that $\epsilon > 0$. The left-hand side of (A10) is increasing in u_t whereas the right-hand side does not depend on u_t . This implies there exists a value of u_t , $u^w(s_{t-1})$ (independent of s_{t-1} by Observation 1) such that $u_t \leq u^w(s_{t-1})$ implies that $s_t^w = 1$ ($=0$). Suppose $u^w(s_{t-1}) \leq \bar{u}_t$ (the argument for the reverse inequality is analogous).

First, consider u_t in the interval $(-\bar{u}_t, \bar{u}_t)$. For all u_t in this interval, $s_t^w = s_t^v = 0$. Varying u_t in this interval gives:

$$W(y_{t-1}, s_{t-1}, u_t) - \beta E_t W(y_{t-1} + \alpha_0, 0, u_{t+1}) = K_2 \quad (A11)$$

where K_2 does not depend upon u_t . As there is no investment we see from Observation 2 that $W(.)$ is independent of u_t in the interval $(-\bar{u}_t, \bar{u}_t)$.

Consider a value of u_t in the interval $(u^w(s_{t-1}), \bar{u}_t)$. This implies $s_t^v = s_t^w = 1$. Considering variations of u_t in this interval gives

$$\mathbb{W}(y_{t-1}, s_{t-1}, u_t) - \beta E_t \mathbb{W}(y_{t-1} + \alpha_0 + \alpha_1 u_t + u_t \rho, v_{t+1}) = K_3 + (\phi_u - \beta \phi_\alpha) u_t \quad (\text{A12})$$

However $\mathbb{W}(\cdot, \cdot, \cdot)$ is linear in u_{t-1} , which is linear in u_{t-1} with coefficient B_u , thus

$$\mathbb{W}(y_{t-1}, s_{t-1}, u_t) = K_4 + \phi_u u_t \quad (\text{A13})$$

Thus over this range too $\mathbb{W}(\cdot, \cdot, \cdot)$ is linear in u_t , with coefficient $B_u = 1/(1-\beta)$.

We also need to consider the interval $(\bar{T}_0 - \bar{T}_1 s_{t-1}, v^w(s_{t-1}))$ in which $s_{t-1}^w = 0$ and \mathbb{W}_t does not depend on u_t over $(\bar{T}_0 - \bar{T}_1 s_{t-1}, u^w(s_{t-1}))$. Thus the coefficient of u_t of the value function $\mathbb{W}(\cdot, \cdot, \cdot)$ has the same form as the value function in (7)-(8).

Finally we need to check how $\mathbb{W}(\cdot)$ is influenced by s_{t-1} . This variable only takes the values 0 and 1. Thus any general function $h(s_{t-1})$ can be written as $k_4 s_{t-1}$, implying any value function that satisfies (7) must have the same form as $V(\cdot)$. QED

References

- Abernathy, W. (1980) *The Productivity Dilemma*, Baltimore John Hopkins
- Acemoglu, D. and Scott, A. (1993) "A theory of economic fluctuations: Increasing Returns and Temporal Agglomeration" Centre For Economic Performance Discussion Paper No.163
- Acemoglu, D. and Scott, A. (1994) "Asymmetries in the cyclical behavior of UK labor markets", *Economic Journal* November 1994
- Arrow, K. (1974) *The Limits of Organization*, New York, Norton.
- Bahk B. H. and M. Gort (1993) "Decomposing learning by doing in new plants" *Journal of Political Economy* Vol 101, pp 561-583.
- Bertola, G. and R. Caballero (1990) "Kinked adjustment costs and aggregate dynamics" *NBER Macroeconomic Annual* MIT Press.
- Bond, S., Elston, J, Mairesse, J and Mulkay, D (1994) "A comparison of empirical investment equations using company panel data for France, Germany, Belgium and the UK", Oxford University mimeo
- Bond, S. and Meghir, C. (1994) "Dynamic investment models and the firm's financial policy", *Review of Economic Studies*, 61, 197-222
- Caballero, R. and Engel, E (1991) "Dynamic (S,s) Economies", *Econometrica*, 59, 1659-1686
- Caballero, R. and Engel, E (1994) "Explaining investment dynamics in US manufacturing: A generalized Ss approach" MIT Mimeo
- Clark, P. K. (1987) "The cyclical component of U.S economic activity", *Quarterly Journal of Economics*, 102, 703-735
- Cooper, R. and Haltiwanger, J (1993) "The aggregate implications of machine replacement: Theory and evidence", *American Economic Review*, 83, 360-82
- Cooper, R, Haltiwanger, J and Power, L (1994) "Machine replacement and the business cycle : Lumps and Bumps", Boston University mimeo
- Cooper, S. (1994) "Multiple regimes in U.S. output fluctuations", Harvard University mimeo
- Davis, S.J and Haltiwanger, J (1992); "Gross Job Creation, Gross Job Destruction, and Employment Reallocation", *Quarterly Journal of Economics* Vol 107, 819-864
- Diebold, F. X. and Rudebusch G. D. (1990) "A non-parametric investigation of duration dependence in the American business cycle", *Journal of Political Economy*, 98, 596-616
- Diebold, F.X and Rudebusch, G.D (1994) "Measuring business cycles: A modern perspective", *Review*

of Economics and Statistics Vol 78, 67-77.

Doms, M. and T. Dunne (1993); "Capital adjustment patterns in manufacturing plants" Mimeo

Durlauf, S. N. (1991) "Multiple equilibria and persistence in aggregate fluctuations", *American Economic Review, papers and proceedings*, 81, 70-74

Durlauf, S. N. (1993) "Nonergodic economic growth", *Review of Economic Studies*, 60, 349-366

Freeman, C. (1982) *The Economics of Industrial Innovation*, Cambridge, MIT Press

Garcia, R (1992) "Asymptotic null distribution of the likelihood ratio test in Markov switching models", CRDE, Montreal, mimeo

Hall, R. E. (1991) *Booms and Recessions in a Noisy Economy*, Yale:Yale University Press

Hamilton, J. D. (1989) "A new approach to the economic analysis of non-stationary time series and the business cycle", *Econometrica*, 57, 357-384

Hansen, B.E (1992) "The likelihood ratio test under nonstandard conditions : testing the Markov switching model of GNP", *Journal of Applied Econometrics*, 7, S61-82

Harvey, A. C. (1985) "Trends and cycles in macroeconomic time series", *Journal of Business and Economic Statistics*, 3, 216-227

Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press

Hirsch, W. Z. (1952) "Manufacturing progress functions" *Review of Economics and Statistics* Vol 34, pp 143-155.

Lahiri, K. and Moore, G. (1991) *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge: Cambridge University Press

Lieberman, M. B. (1984) "The learning curve and pricing curve in chemical processing industries" *Rand Journal of Economics* Vol 15, pp 213-28

Myers, S. and D. G. Marquis (1969) *Successful Industrial Innovations*, Washington DC, National Science Foundation.

Neftci, S. N. (1984) "Are economic time series asymmetric over the business cycle?", *Journal of Political Economy*, 92, 307-28

Nelson, R. R. and S. G. Winter (1982) *An Evolutionary Theory of Economic Change*, Cambridge, MIT Press.

Pennings, J. M. and A. Buitendam (1987); *New Technology as Organizational Innovation; The Development and Diffusion of Microelectronics* eds, Ballinger Publishing Company

Scarf, H. E. (1959) "The optimality of (S,s) policies in the dynamic inventory problem" in K.Arrow, S.Karlin, P.Supes (eds) *Mathematical Methods in Social Sciences*, Stanford University Press

Schankerman, M. (1991); "Revisions of Investment Plans and the Stock Market Rate of Return", NBER Working Paper No.3937

Sichel, D. E. (1993); "Business Cycle Asymmetry: A Deeper Look" *Economic Inquiry* Vol 31, 224-236.
Stock, J. H. and Watson, M. (1989) "New indexes of coincident and leading economic indicators", in Blanchard, O and Fisher, S (eds) *NBER Macroeconomics Annual 1989*

Tushman, M. L. and P. Anderson (1986); "Technological discontinuities and organizational environments" *Administrative Science Quarterly* Vol 31, pp 439-465.

Watson, M. (1986) "Univariate detrending methods with stochastic trends", *Journal of Monetary Economics*, 18, 49-761