Econometrica, Vol. 57, No. 2 (March, 1989), 357–384

A NEW APPROACH TO THE ECONOMIC ANALYSIS OF NONSTATIONARY TIME SERIES AND THE BUSINESS CYCLE

BY JAMES D. HAMILTON[1]

This paper proposes a very tractable approach to modeling changes in regime. The parameters of an autoregression are viewed as the outcome of a discrete-state Markov process. For example, the mean growth rate of a nonstationary series may be subject to occasional, discrete shifts.

The econometrician is presumed not to observe these shifts directly, but instead must draw probabilistic inference about whether and when they may have occurred based on the observed behavior of the series. The paper presents an algorithm for drawing such probabilistic inference in the form of a nonlinear iterative filter. The filter also permits estimation of population parameters by the method of maximum likelihood and provides the foundation for forecasting future values of the series.

An empirical application of this technique to postwar U.S. real GNP suggests that the periodic shift from a positive growth rate to a negative growth rate is a recurrent feature of the U.S. business cycle, and indeed could be used as an objective criterion for defining and measuring economic recessions. The estimated parameter values suggest that a typical economic recession is associated with a 3% permanent drop in the level of GNP.

KEYWORDS: Switching regression, segmentation, nonstationary, business cycle, nonlinear filtering, regime changes.

1. INTRODUCTION AND SUMMARY

A NUMBER OF RECENT STUDIES have sought to characterize the nature of the long term trend in GNP and its relation to the business cycle. Researchers such as Beveridge and Nelson (1981), Nelson and Plosser (1982), and Campbell and Mankiw (1987a, b) explored this question using ARIMA models or ARMA processes around a deterministic trend. Others, such as Harvey (1985), Watson (1986), and Clark (1987) based their analyses on linear unobserved components models. A third approach employs the co-integrated specification of Engle and Granger (1987), whose relevance for business cycle research is examined in a fascinating paper by King, Plosser, Stock, and Watson (1987).

These approaches are based on the assumption that first differences of the log of GNP follow a linear stationary process; that is, in all of the above studies, optimal forecasts of variables are assumed to be a linear function of their lagged values. In this paper I suggest a modest alternative to these currently popular approaches to nonstationarity, exploring the consequences of specifying that first differences of the observed series follow a nonlinear stationary process rather than a linear stationary process. A variety of parameterizations for characterizing nonlinear dynamics have recently been proposed, and there has now accumulated

[1]I am indebted to John Cochrane, Angus Deaton, Robert Engle, Marjorie Flavin, Kevin Hassett, and anonymous referees for comments on earlier drafts of this paper. This material is based upon work supported by the National Science Foundation under Grant No. SES-8720731. The Government has certain rights to this material.

abundant evidence that departures from linearity are an important feature of many key macro series. Studies establishing such nonlinearities include the bispectral analysis of Hinich and Patterson (1985), documentation of business cycle asymmetries by Neftci (1984) and Sichel (1987), the ARCH-M model of Engle, Lilien, and Robins (1987), Stock's (1987) time transformation, chaos models (Brock and Sayers, 1988), Gallant and Tauchen's (1987) "seminonparametric" approach to modeling dynamics, and Quah's (1987) "clinging" process.

The nonlinearities with which my paper is concerned arise if the process is subject to discrete shifts in regime—episodes across which the dynamic behavior of the series is markedly different. My basic approach is to use Goldfeld and Quandt's (1973) Markov switching regression to characterize changes in the parameters of an autoregressive process. For example, the economy may either be in a fast growth or slow growth phase, with the switch between the two governed by the outcome of a Markov process. Building upon ideas developed by Cosslett and Lee (1985), a nonlinear filter and smoother are presented for uncovering optimal statistical estimates of the state of the economy based on observations of output. As in the Kalman filter, one is using the time path of an observed series to draw inference about an unobserved state variable. But whereas the Kalman filter is a linear algorithm for generating estimates of a continuous unobserved state vector, the filter and smoother in this paper provide nonlinear inference about a discrete-valued unobserved state vector.

A very similar stochastic specification has also been explored by Aoki (1967, p. 131), Tong (1983, p. 62), and Sclove (1983), though the statistical approach of these researchers was quite different from the one suggested here. Aoki discussed control of such systems but did not develop the estimation algorithm presented in this paper. Tong treated the shifts in regime as directly observable, whereas the core of my paper addresses optimal probabilistic inference about such shifts based on the observed behavior of GNP. Sclove calculated what the likelihood function would have been if the regimes were observable, and then assumed that the actual historical regimes were those that would make this joint likelihood of GNP along with unobserved regimes as big as possible. My approach, by contrast, is to solve for the actual marginal likelihood function for GNP, maximize this likelihood function with respect to population parameters, and then use these parameters and the data to draw the optimal statistical inference about the unobserved regimes.

My algorithm might also be viewed as formalizing the statistical identification of "turning points" of a time series. Modern treatments by Wecker (1979), Neftci (1982), and Diebold and Rudebusch (1987) provide references to some of the earlier work and interest on this question. Wecker discussed optimal forecasts of an "indicator function" (e.g., $z_t = 1$ if both $y_{t-1} < y_t$ and $y_t > y_{t+1}$). Wecker's indicator is imposed more or less arbitrarily on an otherwise linear process; in my specification, by contrast, the "turning point" is a structural event that is inherent in the data-generating process. Neftci (1982) analyzed the case where (1) only the most recent turning point influences the density function for current observations, and (2) there is known to be a possibility of at most one turning point observed during a given interval (t_1, t_2). These assumptions could also be imposed as a special case of the general framework studied here, generating Neftci's algorithm for dating turning points as a special case of the basic filter used in this study.

The filter also has a clear analog in the analysis of Liptser and Shiryayev (1977), who developed a nonlinear continuous-time filter for a similar problem.[2] The discrete-time filter developed here has three distinct advantages over their treatment. First, if one used Liptser and Shiryayev's formula (which is only strictly valid for continuous time) to approximate discrete changes over short intervals of time, in principle one could end up generating a probability outside the unit interval. By contrast, all probabilities generated by the filter and smoother proposed in this paper are exact, and so lie in $[0,1]$ by construction. Second, a natural byproduct of the discrete-time filter used here is evaluation of the sample likelihood, permitting ready estimation and hypothesis testing about the system's parameters. Third, the specification adopted in this paper fits in neatly as a complement to conventional time series tools and techniques; for example, present value calculations turn out to be quite straightforward.

My approach could also be viewed as a natural extension of Neftci's (1984) analysis of U.S. unemployment data. In Neftci's specification, the economy is said to be in state 1 whenever unemployment is rising and in state 2 whenever unemployment is falling, with transitions between these two states modeled as the outcome of a second-order Markov process. In my paper, by contrast, the unobserved state is only one of many influences governing the dynamic process followed by output, so that even when the economy is in the "fast growth" state, output in principle might be observed to decrease.

The paper applies the technique to postwar U.S. data on real GNP. One possible outcome of maximum likelihood estimation of parameters might have been the identification of long-term trends in the U.S. economy, separating periods with faster growth from those with slower growth. In fact, this is not what was found. Instead, the best empirical fit to the data is obtained when the growth states of the Markov process are associated in a very direct way with the business cycle. A positive growth rate is associated with normal times, and a negative growth rate associated with recessions. Indeed, the best statistical estimates of which quarters were historically characterized by negative growth states for the U.S. economy are remarkably similar to NBER dating of business cycles, and could be used as an alternative objective algorithm for dating business cycles. The results complement the findings by Nelson and Plosser (1982) and Campbell and Mankiw (1987a, b), who concluded that business cycles are associated with a large permanent effect on the long run level of output. The estimates also provide empirical support for the proposition that the dynamics of recessions are qualitatively distinct from those of normal times in a clear statistical sense, and reinforce Neftci's (1984) and Sichel's (1987) evidence on the asymmetry of U.S. business cycles.

[2]See Liptser and Shiryayev (1977, Theorem (9.1), p. 333).

The plan of the paper is as follows. Section 2 specifies the basic model of trend explored in the paper, and compares it with an ARIMA model with normally distributed innovations. Section 3 characterizes the optimal forecast of the future level of a series generated by such a trend. Section 4 presents one example of how this nonlinear trend might interact with a linear process to generate data, and discusses maximum likelihood estimation and inference about the unobserved state for this case. Section 5 applies the technique to postwar U.S. data on real GNP. Section 6 explores the implications for defining and measuring business cycles, and provides a comparison of alternative approaches. Section 7 presents diagnostics comparing the model with the standard ARIMA specification, while Section 8 addresses the long-term consequences of an economic recession. Brief conclusions are offered in Section 9.

2. A MARKOV MODEL OF TREND

Let n_t denote the trend component of a particular time series $\tilde{y}_t$. I will say that n_t obeys a *Markov trend in levels* if

$$n_t = \alpha_1 \cdot s_t + \alpha_0 + n_{t-1} \tag{2.1}$$

where $s_t = 0$ or 1 denotes the unobserved state of the system.[3] I assume that the transition between states is governed by a first-order Markov process:

$$\begin{aligned} &\text{Prob}\,[S_t = 1 \mid S_{t-1} = 1] = p, \\ &\text{Prob}\,[S_t = 0 \mid S_{t-1} = 1] = 1 - p, \\ &\text{Prob}\,[S_t = 0 \mid S_{t-1} = 0] = q, \\ &\text{Prob}\,[S_t = 1 \mid S_{t-1} = 0] = 1 - q. \end{aligned} \tag{2.2}$$

Generalization to a higher-order process and to more than two states is discussed below.

I will describe $\hat{n}_t \equiv \exp(n_t)$ as exhibiting a *Markov trend in logs*.

The stochastic process for S_t (equation 2.2) is strictly stationary, and admits the following AR(1) representation:

$$s_t = (1 - q) + \lambda s_{t-1} + v_t, \tag{2.3}$$

$$\lambda \equiv -1 + p + q, \tag{2.4}$$

where conditional on $S_{t-1} = 1$,

$$\begin{aligned} V_t &= (1 - p) && \text{with probability } p, \\ V_t &= -p && \text{with probability } 1 - p, \end{aligned}$$

conditional on $S_{t-1} = 0$,

$$\begin{aligned} V_t &= -(1 - q) && \text{with probability } q, \\ V_t &= q && \text{with probability } 1 - q. \end{aligned}$$

[3] I adopt the usual notational convention that for discrete-valued variables, capital letters denote the random variable and small letters a particular realization. Both interpretations of course apply to equations such as (2.1), in which I will use small letters by convention.

On the basis of representation (2.3), then, one can view (2.1) as a special case of a standard ARIMA model, albeit with a somewhat unusual probability distribution of the innovation sequence $\{V_t\}$. It is therefore useful to describe in some detail the differences between (2.3) and an AR(1) process driven by normally distributed innovations.

Before doing so, however, note some of the essential properties of (2.3). From (2.3) and the fact that $E_0 V_t = 0$ for all $t > 0$, we see

$$(2.5) \qquad E_0 S_t = \frac{(1-q)(1-\lambda^t)}{(1-\lambda)} + \lambda^t E_0 S_0$$

where E_0 denotes the expectation conditional on information available at date zero (which need not include observation of s_0). Observing that $E_0 S_t$ can be interpreted as the probability that $S_t = 1$ given information available at time zero (denoted $P_0[S_t = 1]$), (2.5) can be rewritten

$$(2.6) \qquad P_0[S_t = 1] = \pi + \lambda^t(\pi_0 - \pi)$$

where

$$(2.7) \qquad \pi \equiv (1-q)/(1-p+1-q),$$
$$\pi_0 \equiv P_0[S_0 = 1].$$

Asymptotically, then, the conditional probability converges to the limiting unconditional probability given by

$$P[S_t = 1] = \pi.$$

As in the case of an ARIMA process with normally distributed innovations, the error term V_t in equation (2.3) is uncorrelated with lagged values of S_t,

$$E\left[V_t | S_{t-j} = 1\right] = E\left[V_t | S_{t-j} = 0\right] = 0 \qquad \text{for} \quad j = 1, 2, \ldots .$$

In contrast to the normal case, however, V_t is not statistically independent of lagged values of S_t, e.g.,

$$E\left[V_t^2 | S_{t-1} = 1\right] = p(1-p),$$
$$E\left[V_t^2 | S_{t-1} = 0\right] = q(1-q).$$

The latter property makes an important difference when noise is added to the system. For example, in the model that I will fit to data, I assume that the state s_t is not observed directly, but instead is one of many factors influencing an observed series. To appreciate the difference that arises in this case between (2.1) and an ARIMA model with normal innovations, consider the simplest possible example:

$$(2.8) \qquad y_t = s_t + \varepsilon_t.$$

Here y_t is a stationary process (perhaps the first difference of $\tilde{y}_t$) and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is an i.i.d. series independent of V_{t-j} for all j. Applying $(1 - \lambda L)$ where L is the lag operator ($L^j x_t = x_{t-j}$) to (2.8),

$$y_t - \lambda y_{t-1} = (1-q) + v_t + \varepsilon_t - \lambda \varepsilon_{t-1}. \tag{2.9}$$

The error term on the right-hand side of (2.9) admits an MA(1) representation,

$$v_t + \varepsilon_t - \lambda \varepsilon_{t-1} = u_t - \theta u_{t-1},$$

or

$$\begin{aligned} u_t = {} & v_t + \theta v_{t-1} + \theta^2 v_{t-2} + \theta^3 v_{t-3} + \cdots + \varepsilon_t + (\theta - \lambda)\varepsilon_{t-1} \\ & + (\theta-\lambda)\theta\varepsilon_{t-2} + (\theta-\lambda)\theta^2\varepsilon_{t-3} + \cdots \end{aligned} \tag{2.10}$$

where θ is the value less than one in absolute value that, along with σ_u^2, satisfies

$$(1+\theta^2)\sigma_u^2 = (1+\lambda^2)\sigma_\varepsilon^2 + \sigma_v^2 \tag{2.11}$$

$$-\theta\sigma_u^2 = -\lambda\sigma_\varepsilon^2, \tag{2.12}$$

where

$$\begin{aligned} \sigma_v^2 &= E(V_t^2) \\ &= p(1-p)\pi + q(1-q)(1-\pi). \end{aligned} \tag{2.13}$$

As in the case of V_t, the innovation U_t is uncorrelated with U_{t-j} for $j > 0$, but is not independent. An earlier version of this paper illustrated the relevance of this point by way of example, showing that while $E[U_t(U_{t-1} - \theta U_{t-2})] = 0$, it nonetheless is the case that

$$E\left[U_t(U_{t-1} - \theta U_{t-2})^2\right] = \theta\left(\frac{(1-p)(1-q)(p-q)}{(1-\lambda)}\right)\left(\frac{\theta\lambda^2 - 2\lambda - 1}{1-\theta\lambda}\right),$$

in general not zero. What this means in practical terms is that while one could use the ARMA(1,1) representation

$$y_t - \lambda y_{t-1} = (1-q) + u_t - \theta u_{t-1}$$

as a basis for forecasting y_{t+j} as a linear function of $y_t, y_{t-1}, \ldots,$ these forecasts are not optimal; nonlinear forecasts that exploit the serial dependence of the white noise series U_t are superior.[4] From (2.6), these optimal forecasts are given by

$$E_t y_{t+j} = \pi + \lambda^j \cdot \{P[S_t = 1 | y_t, y_{t-1}, \ldots] - \pi\}$$

where $P[S_t = 1 | y_t, y_{t-1}, \ldots]$ is the nonlinear function of $y_t, y_{t-1}, \ldots$ to be presented in Section 4.

Thus, the essential differences between the specification (2.1) and a standard ARIMA model with normal innovations are twofold. First, (2.1) specifies that the growth rate $n_t - n_{t-1}$ need not change every period, but rather only does so in response to occasional, discrete events. Second, when added to a linear normal process, (2.1) generates a nonlinear process for the observed series for which,

[4] See Granger (1983) on this general issue.

while an ARIMA representation exists, it does not generate optimal forecasts of the future value of the series.

3. FORECASTING AND PRESENT VALUE CALCULATIONS

3.1. *Markov Trend in Levels*

Let i_t denote the cumulative number of "ones" since time zero,

$$i_t \equiv s_1 + s_2 + \cdots + s_t,$$

so from (2.1),

$$n_t = n_0 + \alpha_1 i_t + \alpha_0 t. \tag{3.1}$$

Recall from (2.6) that

$$E\{S_t | \text{Prob}[S_0 = 1] = \pi_0\} = \pi + \lambda^t(\pi_0 - \pi) \tag{3.2}$$

and so from (3.1),

$$\begin{aligned} E_0\{N_t | E_0[N_0] = n_0, \text{Prob}[S_0 = 1] = \pi_0\} &= n_0 + \alpha_1\left[\pi t + \sum_{\tau=1}^{t} \lambda^\tau(\pi_0 - \pi)\right] + \alpha_0 t \\ &= n_0 + [\alpha_1\pi + \alpha_0]t + [\alpha_1\lambda(1 - \lambda^t)/(1 - \lambda)][\pi_0 - \pi]. \end{aligned} \tag{3.3}$$

The limiting growth rate as $t \to \infty$ is seen from (3.3) to be independent of information about the state of the system at date 0:

$$\lim_{t \to \infty} E[(N_{t+1} - N_t) | n_0, \pi_0] = \alpha_1\pi + \alpha_0.$$

Intuitively, we know from equation (2.6) that for large t the economy will be in state 1 with probability π, in which case the growth rate would be $\alpha_1 + \alpha_0$, whereas the economy will be in state 0 with probability $1 - \pi$, in which case the growth rate would be α_0; hence the expected growth rate is $\alpha_1\pi + \alpha_0$. Furthermore, if one had no useful information about the state of the system at date 0, $\pi_0 = \pi$ and (3.3) implies that this limiting growth rate would be the basis for constructing forecasts of N_t for all finite t. On the other hand, if one did have useful information that, say, $\pi_0 > \pi$, then for $\alpha_1\lambda > 0$, $E[N_t | P_0(S_0 = 1) = \pi_0]$ would be systematically larger than $E[N_t | P_0(S_0 = 1) = \pi]$ for all t, with the difference growing with t as the term $(1 - \lambda^t)$ goes to unity. In particular, if we compare certain knowledge that $S_0 = 1$ $(\pi_0 = 1)$ with certain knowledge that $S_0 = 0$ $(\pi_0 = 0)$, we see

$$\lim_{t \to \infty} \{E[N_t | S_0 = 1] - E[N_t | S_0 = 0]\} = \alpha_1\lambda/(1 - \lambda). \tag{3.4}$$

So, while information about the state of the economy at date 0 has no effect on the long run growth rate $(N_{t+1} - N_t)$, it does exert a permanent effect on the level N_t.[5]

[5] An analogous result of course characterizes a standard ARIMA$(p, 1, q)$ process. See Beveridge and Nelson (1981, p. 155).

The discounted present value can also be evaluated from (3.3):

$$E\left\{\sum_{t=0}^{\infty}\beta^t N_t \mid n_0, \pi_0\right\} = \frac{n_0}{(1-\beta)} + \alpha_1\left[\frac{\beta(1-q)}{(1-\beta)^2(1-\beta\lambda)} + \frac{\beta\lambda\pi_0}{(1-\beta)(1-\beta\lambda)}\right] + \frac{\alpha_0\beta}{(1-\beta)^2}. \tag{3.5}$$

3.2. *Markov Trend in Logs*

Here I characterize forecasts of future values of a series that follows a Markov trend in logs by exploiting a simple vector recursion in expected values.

Let $P_\tau[A, B]$ denote the probability that events A and B will occur together, conditional on information available at τ. Note that the following recursion,

$$P_0[I_t = i, S_t = 1] = p \cdot P_0[I_{t-1} = i-1, S_{t-1} = 1] + (1-q)\cdot P_0[I_{t-1} = i-1, S_{t-1} = 0], \tag{3.6}$$

holds for $t = 1, 2, \ldots$ and $i = 1, 2, \ldots, t$. For $i = 0$ we of course have

$$P_0[I_t = 0, S_t = 1] = 0 \tag{3.7}$$

holding for $t = 1, 2, \ldots$. Similarly, the recursion

$$P_0[I_t = i, S_t = 0] = (1-p)\cdot P_0[I_{t-1} = i, S_{t-1} = 1] + q \cdot P_0[I_{t-1} = i, S_{t-1} = 0], \tag{3.8}$$

holds for $t = 1, 2, \ldots$ and $i = 0, 1, \ldots, t-1$, with

$$P_0[I_t = t, S_t = 0] = 0 \tag{3.9}$$

for $t = 1, 2, \ldots$.

Let $\hat{\alpha}_1 \equiv \exp(\alpha_1)$ and $\hat{\alpha}_0 \equiv \exp(\alpha_0)$. Multiplying equation (3.6) by $\hat{\alpha}_1^i \hat{\alpha}_0^t$, summing for $i = 1, 2, \ldots, t$, and using (3.7) yields

$$\begin{aligned}\sum_{i=0}^{t} \hat{\alpha}_1^i \hat{\alpha}_0^t \cdot P_0[I_t = i, S_t = 1] &= [\hat{\alpha}_1 \hat{\alpha}_0 p] \cdot \sum_{j=0}^{t-1} \hat{\alpha}_1^j \hat{\alpha}_0^{t-1} \cdot P_0[I_{t-1} = j, S_{t-1} = 1] \\ &\quad + [\hat{\alpha}_1 \hat{\alpha}_0 (1-q)] \cdot \sum_{j=0}^{t-1} \hat{\alpha}_1^j \hat{\alpha}_0^{t-1} \cdot P_0[I_{t-1} = j, S_{t-1} = 0].\end{aligned} \tag{3.10}$$

Similarly, multiplying (3.8) by $\hat{\alpha}_1^i \hat{\alpha}_0^t$, summing for $i = 0, 1, \ldots, t-1$, and using

(3.9) gives

$$\text{(3.11)} \quad \sum_{i=0}^{t} \hat{\alpha}_1^i \hat{\alpha}_0^t \cdot P_0[I_t = i,\ S_t = 0]$$
$$= [(1-p)\hat{\alpha}_0] \cdot \sum_{j=0}^{t-1} \hat{\alpha}_1^j \hat{\alpha}_0^{t-1} \cdot P_0[I_{t-1} = j,\ S_{t-1} = 1]$$
$$+ [q\hat{\alpha}_0] \cdot \sum_{j=0}^{t-1} \hat{\alpha}_1^j \hat{\alpha}_0^{t-1} \cdot P_0[I_{t-1} = j,\ S_{t-1} = 0].$$

Define

$$\text{(3.12)} \quad M_0(t,s) = \sum_{i=0}^{t} \hat{\alpha}_1^i \hat{\alpha}_0^t \cdot P_0[I_t = i,\ S_t = s]$$

for $s = 0, 1$ and write (3.10) and (3.11) as

$$\begin{bmatrix} M_0(t,1) \\ M_0(t,0) \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_0\hat{\alpha}_1 p & \hat{\alpha}_0\hat{\alpha}_1(1-q) \\ \hat{\alpha}_0(1-p) & \hat{\alpha}_0 q \end{bmatrix} \begin{bmatrix} M_0(t-1,1) \\ M_0(t-1,0) \end{bmatrix},$$

or, defining

$$B \equiv \begin{bmatrix} \hat{\alpha}_1 p & \hat{\alpha}_1(1-q) \\ (1-p) & q \end{bmatrix},$$

we have

$$\text{(3.13)} \quad \begin{bmatrix} M_0(t,1) \\ M_0(t,0) \end{bmatrix} = \hat{\alpha}_0 B \begin{bmatrix} M_0(t-1,1) \\ M_0(t-1,0) \end{bmatrix}.$$

Note from (3.12) that $M_0(0, s) = P_0[S_0 = s]$. Thus (3.13) has the solution

$$\begin{bmatrix} M_0(t,1) \\ M_0(t,0) \end{bmatrix} = \hat{\alpha}_0^t B^t \begin{bmatrix} \pi_0 \\ 1-\pi_0 \end{bmatrix}.$$

Solving for the roots of $|\mu I - B| = 0$, we see

$$\mu_1 + \mu_2 = q + p\hat{\alpha}_1,$$
$$\mu_1\mu_2 = \hat{\alpha}_1(-1 + p + q).$$

Following Chiang (1980, pp. 148–152), write

$$\text{(3.14)} \quad B^t = T \begin{bmatrix} \mu_1^t & 0 \\ 0 & \mu_2^t \end{bmatrix} T^{-1}$$

where

$$T = \begin{bmatrix} (\mu_1 - q) & (\mu_2 - q) \\ (1-p) & (1-p) \end{bmatrix},$$

$$T^{-1} = \frac{1}{(\mu_1 - \mu_2)(1-p)} \begin{bmatrix} (1-p) & (q-\mu_2) \\ -(1-p) & (\mu_1 - q) \end{bmatrix}.$$

The expected value of the level of a series that follows a Markov trend in logs is then seen to be

$$(3.15)\qquad E_0\hat{N}_t = \hat{n}_0[M_0(t,1)+M_0(t,0)]$$
$$= \hat{n}_0\cdot[1\quad 1]\hat{\alpha}_0^t B^t[\pi_0\quad 1-\pi_0]'$$
$$= \frac{\hat{n}_0\hat{\alpha}_0^t\{(k_0-\mu_2)\mu_2^t-(k_0-\mu_1)\mu_1^t\}}{(\mu_1-\mu_2)}$$

where

$$k_0 \equiv [\mu_1\mu_2/\hat{\alpha}_1][\pi_0+\hat{\alpha}_1(1-\pi_0)]$$
$$=[-1+p+q][\pi_0+\hat{\alpha}_1(1-\pi_0)].$$

Normalizing $\mu_1>\mu_2$, we see that, as in the case of a Markov trend in levels, the long-run growth rate is independent of information about the initial state:

$$\lim_{t\to\infty}\frac{E_0\hat{N}_{t+1}}{E_0\hat{N}_t}=\hat{\alpha}_0\mu_1$$

but a change in the current state exerts a permanent effect on the future level of the series,

$$(3.16)\qquad \lim_{t\to\infty}\frac{E_0\{\hat{N}_t|\pi_0=1\}}{E_0\{\hat{N}_t|\pi_0=0\}}=\frac{\mu_1-(-1+p+q)}{\mu_1-\hat{\alpha}_1(-1+p+q)}.$$

From (3.15), the present value is

$$(3.17)\qquad E_0\sum_{t=0}^{\infty}\beta^t\hat{N}_t=\frac{\hat{n}_0(1-k_0\beta\hat{\alpha}_0)}{1-\beta\hat{\alpha}_0(p\hat{\alpha}_1+q)+\beta^2\hat{\alpha}_0^2(-1+p+q)\hat{\alpha}_1}.$$

4. ESTIMATION, FILTERING, AND SMOOTHING

4.1. *Stochastic Specification*

Several options are available for combining the trend term n_t with another stochastic process. Here I discuss the approach that results in the computationally simplest maximum likelihood estimation.

Suppose we have observations on a time series $\{\tilde{y}_t\}$. Specify

$$(4.1)\qquad \tilde{y}_t = n_t+\tilde{z}_t$$

where n_t is as given in (2.1) and (2.2) and $\tilde{z}_t$ follows a zero mean ARIMA$(r,1,0)$ process:

$$(4.2)\qquad \tilde{z}_t-\tilde{z}_{t-1}=\phi_1(\tilde{z}_{t-1}-\tilde{z}_{t-2})+\phi_2(\tilde{z}_{t-2}-\tilde{z}_{t-3})+\cdots$$
$$+\phi_r(\tilde{z}_{t-r}-\tilde{z}_{t-r-1})+\varepsilon_t.$$

I take $\{\varepsilon_t\}$ to be an i.i.d. $N(0, \sigma^2)$ sequence that is independent of $\{n_{t+j}\}$ for all j. Differencing (4.1) and rewriting (4.2) we obtain

$$(4.3)\qquad \begin{aligned} y_t &= \alpha_1 s_t + \alpha_0 + z_t, \\ z_t &= \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_r z_{t-r} + \varepsilon_t, \end{aligned}$$

where $y_t \equiv \tilde{y}_t - \tilde{y}_{t-1}$ and $z_t \equiv \tilde{z}_t - \tilde{z}_{t-1}$.

The econometrician is presumed to observe y_t but not z_t or s_t. I first discuss a filter whereby the econometrician can draw probabilistic inference about the unobserved state s_t given observations on y_t, and then show how evaluation of the sample likelihood is a natural byproduct of the filter. The analysis is closely related to the discussion by Cosslett and Lee (1985), who derived a recursion to evaluate the likelihood function for the case where (4.3) is a standard stationary regression equation with no lagged dependent variables.

4.2. *Filtering*

The *basic filter* accepts as input the joint conditional probability

$$P[S_{t-1} = s_{t-1}, S_{t-2} = s_{t-2}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}]$$

and has as output

$$P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r+1} = s_{t-r+1} | y_t, y_{t-1}, \ldots, y_{-r+1}]$$

along with, as a byproduct, the conditional likelihood of y_t:

$$f(y_t | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}).$$

Note well the notation: $[s_t, s_{t-1}, \ldots, s_{t-r+1}]$ refers to the r most recent values of s whereas $[y_t, y_{t-1}, \ldots, y_{-r+1}]$ denotes the complete history of y observed through date t. By "$P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r+1} = s_{t-r+1} | y_t, y_{t-1}, \ldots, y_{-r+1}]$" I refer to a vector consisting of 2^r elements. For example, suppose $r = 4$. The element indexed by (1, 0, 1, 1) denotes the probability that $S_{t-1} = 1$, $S_{t-2} = 0$, $S_{t-3} = 1$, and $S_{t-4} = 1$. These 16 probabilities sum to unity by construction, and represent an inference about the unobserved state $(s_{t-1}, s_{t-2}, s_{t-3}, s_{t-4})$ based on observations of y through date $t-1$. The algorithm is as follows.

STEP 1: Calculate

$$\begin{aligned} P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}] \\ = P[S_t = s_t | S_{t-1} = s_{t-1}] \times P[S_{t-1} = s_{t-1}, S_{t-2} = s_{t-2}, \ldots, \\ S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}] \end{aligned}$$

where $P[S_t = s_t | S_{t-1} = s_{t-1}]$ is given by (2.2). (Note $P[S_t = s_t | S_{t-1} = s_{t-1}] = P[S_t = s_t | S_{t-1} = s_{t-1}, S_{t-2} = s_{t-2}, \ldots, S_{t-r} = s_{t-r}, y_{t-1}, y_{t-2}, \ldots, y_{-r+1}]$ by the independence and first-order Markov assumptions.)

STEP 2: Calculate the joint conditional density-distribution of y_t and $(S_t, S_{t-1}, \ldots, S_{t-r})$:

$$f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1})$$
$$= f(y_t | S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r}, y_{t-1}, y_{t-2}, \ldots, y_{-r+1})$$
$$\times P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}]$$

where we know

$$f(y_t | S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r}, y_{t-1}, y_{t-2}, \ldots, y_{-r+1})$$
$$= \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2\sigma^2} \left((y_t - \alpha_1 s_t - \alpha_0) - \phi_1 (y_{t-1} - \alpha_1 s_{t-1} - \alpha_0) - \cdots - \phi_r (y_{t-r} - \alpha_1 s_{t-r} - \alpha_0) \right)^2 \right].$$

STEP 3: We then have

$$f(y_t | y_{t-1}, y_{t-2}, \ldots, y_{-r+1})$$
$$= \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} \cdots \sum_{s_{t-r}=0}^{1} f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1})$$

STEP 4: Thus

$$P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_t, y_{t-1}, \ldots, y_{-r+1}]$$
$$= \frac{f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1})}{f(y_t | y_{t-1}, y_{t-2}, \ldots, y_{-r+1})}.$$

STEP 5: The desired output is then obtained from

$$P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r+1} = s_{t-r+1} | y_t, y_{t-1}, \ldots, y_{-r+1}]$$
$$= \sum_{s_{t-r}=0}^{1} P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_t, y_{t-1}, \ldots, y_{-r+1}].$$

One could start up the algorithm with

$$P[S_0 = s_0, S_{-1} = s_{-1}, \ldots, S_{-r+1} = s_{-r+1} | y_0, y_{-1}, \ldots, y_{-r+1}]$$

though evaluating this expression proves to be somewhat involved computationally. I have instead in this paper adopted the simpler expedient of starting the filter with the unconditional probability $P[S_0 = s_0, \; S_{-1} = s_{-1}, \ldots, S_{-r+1} = s_{-r+1}]$, evaluated as follows. Set $P[S_{-r+1} = 1]$ equal to the limiting probability π of the Markov process from equation (2.7), and of course set $P[S_{-r+1} = 0] =$

$1-\pi$. Then for $\tau = -r+2, -r+3, \ldots, 0$ calculate

$$\begin{aligned} P[S_\tau = s_\tau,\ S_{\tau-1} = s_{\tau-1}, \ldots, S_{-r+1} = s_{-r+1}] \\ = P[S_\tau = s_\tau | S_{\tau-1} = s_{\tau-1}] \\ \times P[S_{\tau-1} = s_{\tau-1},\ S_{\tau-2} = s_{\tau-2}, \ldots, S_{-r+1} = s_{-r+1}]. \end{aligned}$$

The final product of this subiteration,

$$P[S_0 = s_0,\ S_{-1} = s_{-1}, \ldots, S_{-r+1} = s_{-r+1}],$$

is then used as input for the basic filter for $t = 1$. The iteration on the basic filter is then repeated for $t = 1, 2, \ldots, T$.

For some applications, one might want to allow the possibility of a permanent change in regime (e.g., $q = 1$). For such applications, we should not set $P[S_{-r+1} = 1]$ from equation (2.7), but should instead treat it as a separate parameter (say π_{-r+1}) to be estimated along with the others.

It is easy to verify that the output of the filter is always a well-defined probability distribution with the terms nonnegative and summing to unity.

Neftci's (1982) algorithm for dating business cycle turning points can be obtained as a special case of the basic filter by setting $q = 1$ and $r = 0$.

One byproduct of the filter is evaluation of the conditional likelihood in Step 3. The sample conditional log likelihood is

$$\begin{aligned} \log f(y_T, y_{T-1}, \ldots, y_1 | y_0, y_{-1}, \ldots, y_{-r+1}) \\ = \sum_{t=1}^{T} \log f(y_t | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}) \end{aligned}$$

which can be maximized numerically with respect to the unknown parameters $(\alpha_1, \alpha_0, p, q, \sigma, \phi_1, \phi_2, \ldots, \phi_r)$, and optionally π_{-r+1} as described above. Obviously the model is unidentified in the sense that the decision of which state to call state 0 and which to call state 1 is arbitrary. I normalize by letting state 1 be the fast growth state and state 0 be the slow growth state, achieved by setting $\alpha_1 + \alpha_0 > \alpha_0$ or $\alpha_1 > 0$.

The logic of the filter is equally valid under much more general specifications. With n rather than 2 states, the input to the filter is a vector consisting of n^r elements, and the summations in Steps 3 and 5 are over $(0, n-1)$ rather than $(0, 1)$. The autoregressive parameters (ϕ) can also be made a function of the regime by replacing ϕ_j in Step 2 with $\phi_j(S_t)$ or $\phi_j(S_{t-j})$. In my (1988) paper I applied the algorithm with the standard deviation $\sigma(S_t)$ also a function of the regime, and extended the estimation theory to a multivariate context where the econometrician wishes to impose the cross-equation restrictions implied by rational expectations. Higher-order dynamics for the regime shift are also conceptually straight-forward—e.g., replace $P[S_t = s_t | S_{t-1} = s_{t-1}]$ in Step 2 with $P[S_t = s_t | S_{t-1} = s_{t-1},\ S_{t-2} = s_{t-2}]$. That is, instead of multiplying each of the 16 numbers in the input to the filter by $p, q, 1-p$, or $1-q$ (depending on the value of s_t and s_{t-1}) one multiplies by one of $p_{11}, p_{12}, \ldots$ depending on the value of s_t, s_{t-1}, and s_{t-2}.

Such extensions are in principle straight-forward. Any problems are chiefly numerical. Identification of the parameters characterizing the dynamics of S_t (p, q, and α_1) separately from those of the Gaussian component ($\phi_1, \phi_2, \ldots, \phi_r$) depends on nonlinearities in the data. There is a practical limit on how complicated we can permit the dynamics for both the regime shift and the Gaussian component to become and still have hope of obtaining useful results.

The relation between my approach and that of Sclove (1983) should now be stated more precisely. Let $\boldsymbol{y} \equiv (y_1, \ldots, y_T)'$, $\boldsymbol{s} \equiv (s_1, \ldots, s_T)'$, and $\boldsymbol{\theta} = (\alpha_1, \alpha_0, p, q, \sigma, \phi_1, \phi_2, \ldots, \phi_r)'$. My filter evaluates $f(\boldsymbol{y}|\boldsymbol{\theta}, y_{-r+1}, \ldots, y_0)$ and maximizes with respect to $\boldsymbol{\theta}$. The MLE $\hat{\boldsymbol{\theta}}$ is then used in a final pass through the filter to draw probabilistic inference about $\boldsymbol{s}$. Sclove, by contrast, would calculate $f(\boldsymbol{y}, \boldsymbol{s}|\boldsymbol{\theta}, y_{-r+1}, \ldots, y_0)$ and maximize with respect to both $\boldsymbol{\theta}$ and $\boldsymbol{s}$. Thus the output of my algorithm is a sequence of conditional probabilities, and the output of Sclove's maximization is an imputed historical sequence for $\boldsymbol{s}$. Sclove's empirical application also opted for the other end of the trade-off between a rich parameterization of the dynamics for the Gaussian component and that for the Markov component. He assumed no autocorrelation for the Gaussian component, whereas I allow four lags; Sclove tested for up to nine different regimes, whereas I permit only two.

4.3. *Smoothing*

Another byproduct of the basic filter is inference about the state s_t based on currently available information,

$$P[S_t = s_t | y_t, y_{t-1}, \ldots, y_{-r+1}] = \sum_{s_{t-1}=0}^{1} \sum_{s_{t-2}=0}^{1} \cdots \sum_{s_{t-r+1}=0}^{1} P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r+1} = s_{t-r+1} | y_t, y_{t-1}, \ldots, y_{-r+1}].$$

Alternatively, one can obtain a more reliable inference about the lagged value of the state using currently available information. For example, using the output from Step 4 of the basic filter, one can calculate an *r-lag smoother*:

$$P[S_{t-r} = s_{t-r} | y_t, y_{t-1}, \ldots, y_{-r+1}] = \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} \cdots \sum_{s_{t-r+1}=0}^{1} P[S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-r} = s_{t-r} | y_t, y_{t-1}, \ldots, y_{-r+1}].$$

A full-sample smoother can be obtained from adapting a suggestion made by Cosslett and Lee (1985) in a slightly different context. Suppose that instead of using

$$P[S_{t-1} = s_{t-1}, S_{t-2} = s_{t-2}, \ldots, S_{t-r} = s_{t-r} | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}]$$

as input into the basic filter, we used in its place

$$P[S_{t-1}=s_{t-1},\ S_{t-2}=s_{t-2},\dots,S_{t-r}=s_{t-r}|S_\tau=\hat{s}_\tau,\\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{t-1},\ y_{t-2},\dots,y_{-r+1}]$$

for some $\tau \leqslant t-1$ and for some choice of $(\hat{s}_\tau, \hat{s}_{\tau-1},\dots,\hat{s}_{\tau-r+1})$ to be specified shortly. Running through the steps of the basic filter, it is easy to verify that the output of the filter would in this case be

$$P[S_t=s_t,\ S_{t-1}=s_{t-1},\dots,S_{t-r+1}=s_{t-r+1}|S_\tau=\hat{s}_\tau,\\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_t,\ y_{t-1},\dots,y_{-r+1}]$$

with byproduct

$$f(y_t|S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{t-1},\ y_{t-2},\dots,y_{-r+1}).$$

Bearing this in mind, the *full-sample smoother* can be obtained as follows.

STEP 1: Run through the basic filter for $t=1,\dots,T$ and store the resulting sequences $P[S_\tau=s_\tau,\ S_{\tau-1}=s_{\tau-1},\dots,S_{\tau-r+1}=s_{\tau-r+1}|y_\tau,\ y_{\tau-1},\dots,y_{-r+1}]$ and $f(y_\tau|y_{\tau-1},\ y_{\tau-2},\dots,y_{-r+1})$ for $\tau=1,2,\dots,T$.

STEP 2: For each τ and for each possible value of the vector $(\hat{s}_\tau, \hat{s}_{\tau-1},\dots,\hat{s}_{\tau-r+1})$, repeat the following:

(a) Set

$$P[S_\tau=s_\tau,\ S_{\tau-1}=s_{\tau-1},\dots,S_{\tau-r+1}=s_{\tau-r+1}|S_\tau=\hat{s}_\tau,\\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{t-1},\ y_{t-2},\dots,y_{-r+1}] \tag{4.4}$$

equal to unity if $s_\tau=\hat{s}_\tau,\ s_{\tau-1}=\hat{s}_{\tau-1},\dots,s_{\tau-r+1}=\hat{s}_{\tau-r+1}$ and zero otherwise.

(b) Repeat the basic filter using (4.4) to start the iteration and iterate over $t=\tau+1,\tau+2,\dots,T$, storing the output from Step 3 of the basic filter as $f(y_t|S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{t-1},\ y_{t-2},\dots,y_{-r+1})$.

(c) The smoothed probabilities are given by

$$\begin{aligned}
&P[S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1}|y_T,\ y_{T-1},\dots,y_{-r+1}]\\
&\quad=P[S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1}|y_\tau,\ y_{\tau-1},\dots,y_{-r+1}]\\
&\quad\times\frac{f(y_{\tau+1}|S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_\tau,\ y_{\tau-1},\dots,y_{-r+1})}{f(y_{\tau+1}|y_\tau,\ y_{\tau-1},\dots,y_{-r+1})}\\
&\quad\times\frac{f(y_{\tau+2}|S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{\tau+1},\ y_\tau,\dots,y_{-r+1})}{f(y_{\tau+2}|y_{\tau+1},\ y_\tau,\dots,y_{-r+1})}\\
&\quad\times\cdots\times\frac{f(y_T|S_\tau=\hat{s}_\tau,\ S_{\tau-1}=\hat{s}_{\tau-1},\dots,S_{\tau-r+1}=\hat{s}_{\tau-r+1},\ y_{T-1},\ y_{T-2},\dots,y_{-r+1})}{f(y_T|y_{T-1},\ y_{T-2},\dots,y_{-r+1})}.
\end{aligned}$$

5. MAXIMUM LIKELIHOOD ESTIMATES FOR U.S. GNP DATA

The above technique was applied to U.S. postwar data on real GNP. The variable used for y_t was 100 times the change in the log of real GNP for

TABLE I

MAXIMUM LIKELIHOOD ESTIMATES OF PARAMETERS AND ASYMPTOTIC STANDARD ERRORS BASED ON DATA FOR U.S. REAL GNP, $t = 1952:\text{II}$ TO $1984:\text{IV}$

Parameter	Estimate	Standard error
α_1	1.522	0.2636
α_0	−0.3577	0.2651
p	0.9049	0.03740
q	0.7550	0.09656
σ	0.7690	0.06676
ϕ_1	0.014	0.120
σ_2	−0.058	0.137
ϕ_3	−0.247	0.107
ϕ_4	−0.213	0.110

$t = 1951:\text{II}$ to $1984:\text{IV}$.[6] Numerical maximization of the conditional log likelihood function led to the maximum likelihood estimates reported in Table I. Also reported are asymptotic standard errors.[7]

One possible outcome that might have been expected a priori would associate the states $s_t = 0$ and 1 with slow and fast growth rates for the U.S. economy, corresponding to decade-long changes in trends. In fact, however, the sample likelihood is maximized by a negative growth rate of -0.4% per quarter during state 0 and a positive growth of $(\alpha_0 + \alpha_1) = +1.2\%$ during state 1. These values clearly correspond to the dynamics of business cycles as opposed to long-term variations in secular growth rates. Indeed, the first- and second-order serial correlation in logarithmic changes of real GNP seem to be better captured by shifts between states rather than by the leading autoregressive coefficients, as indicated by the fact that $\hat{\phi}_1$ and $\hat{\phi}_2$ come out remarkably close to zero. Negative coefficients at lags 3 and 4 suggest the possibility that the method used by the Bureau of Economic Analysis for deseasonalizing introduces spurious periodicity when applied to data generated by a nonlinear process such as this one. These coefficients further suggest that investigating a higher-order Markov process for the trend might also be a fruitful topic for future research.

Figure 1 reports the estimated probability that the economy is in the negative growth state ($P[S_t = 0]$) based on currently available information (panel A) and information available one year later (panel B). A full sample smoother (not shown) was also calculated. The probabilities from the full-sample smoother differed very little from those of the four-lag smoother in panel B. The average absolute difference between these two smoothed series was .016, with the maximum difference occurring in the second quarter of 1956; (the four-lag smoother

[6] The level of GNP is measured at an annual rate in 1982 dollars. Data are from *Business Conditions Digest*, February, 1986, p. 102, Series 50. The order of lags r was set arbitrarily to 4; the basic filter was thus started for $t = 1952:\text{II}$.

[7] Maximization was achieved by a Davidon-Fletcher-Powell routine. Convergence to the global maximum reported in Table I proved relatively robust with respect to a broad range of start-up values. Second derivatives of the log likelihood were calculated numerically, from which asymptotic standard errors were constructed. I would like to thank Kent Wall for use of his DFP algorithm and Steve Stern for use of his second-derivative program.

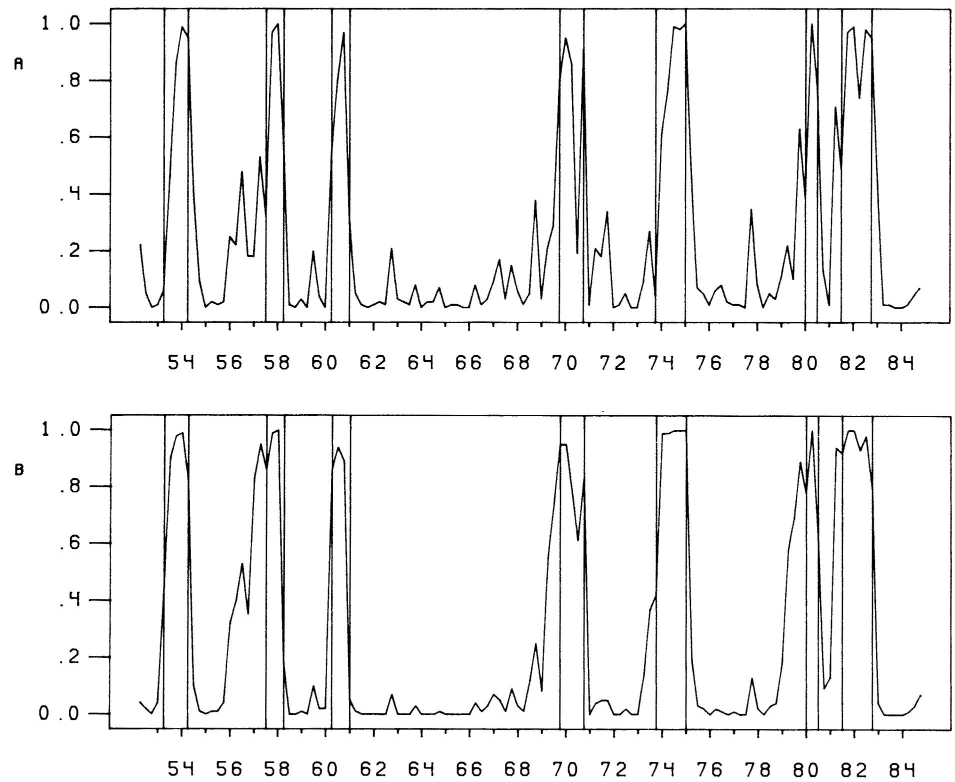


FIGURE 1.—Inferred probability that $S_t = 0$.
Panel (A) reports the inferred probability that the economy was in the falling GNP state at date t using information available at the time ($P[S_t = 0 | y_t, y_{t-1}, \ldots]$). Panel (B) reports the inferred probability that the economy was in the falling GNP state at date t using information available 4 quarters later ($P[S_t = 0 | y_{t+4}, y_{t+3}, \ldots]$).

puts the probability of contraction at .40 for this quarter, whereas the full-sample inference was .15). This suggests that reasonably precise estimates are available from the four-lag smoother associated with the basic filter itself, and it may be unnecessary to employ the full-sample smoother for many applications. Another reasonable alternative to the full-sample smoother is to augment the basic filter with a few additional lags on s.

6. ESTABLISHING THE DATES OF HISTORICAL BUSINESS CYCLES

The specific inferences about the historical incidence of growth states generated by the filter and smoother correspond extremely closely to conventional dating of business cycles, and indeed could be employed as an independent objective algorithm for generating such dating. A sensible metric might be based on whether the econometrician would conclude that the economy is more likely

TABLE II

ALTERNATIVE DATING OF U.S. BUSINESS CYCLE PEAKS AND TROUGHS AS DETERMINED BY (1) NBER, AND (2) PROBABILITY OF BEING IN RECESSION GREATER THAN 0.5 AS DETERMINED FROM FULL-SAMPLE SMOOTHER

NBER		Smoother	
Peak	Trough	Peak	Trough
1953 : III	1954 : II	1953 : III	1954 : II
1957 : III	1958 : II	1957 : I	1958 : I
1960 : II	1961 : I	1960 : II	1960 : IV
1969 : IV	1970 : IV	1969 : III	1970 : IV
1973 : IV	1975 : I	1974 : I	1975 : I
1980 : I	1980 : III	1979 : II	1980 : III
1981 : III	1982 : IV	1981 : II	1982 : IV

than not to be in a recession ($P[S_t = 0 | y_T, y_{T-1}, \ldots, y_{-r+1}] > 0.5$). Dates for postwar business cycles based on this measure are compared with NBER values in Table II.[8] In contrast to NBER dates, my series indicates that the recessions of 1957–58 and 1979–80 immediately followed the oil price increases of 1957 : I associated with the Suez Crisis and 1979 : II associated with the Iranian revolution, respectively.[9] For the other recessions, the two dating techniques are always within three months of each other.

Note that the particular decision rule $P[S_t = 0] > 0.5$ seems to be largely irrelevant for these data. Very few of the smoothed probabilities in panel B of Figure 1 lie between 0.3 and 0.7. The algorithm is usually arriving at a fairly strong conclusion about whether the economy is in a recession. The implicit histogram would also seem to suggest that the filter is not simply fitting parameters to an arbitrary nonlinear process, but rather reflects an underlying pattern in the data of dichotomous shifts between the expansion and contraction phase.

Another interesting implication of the Markov framework is that one can calculate from the maximum likelihood parameter estimates the expected duration of a typical recession and compare this predicted magnitude with the historical average. Conditional on being in state 0, the expected duration of a recession is

$$\sum_{k=1}^{\infty} k q^{k-1}(1-q) = (1-q)^{-1}$$

or 4.1 quarters. The historical average duration of a recession was 4.7 quarters during the postwar period according to the NBER figures. The expected duration of an expansion is likewise $(1-p)^{-1}$ or 10.5 quarters, compared with an average of 14.3 quarters in NBER dating.

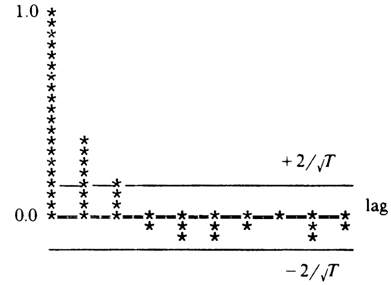
[8] NBER business cycle dates are reported in *Business Conditions Digest* published by the Department of Commerce.

[9] My (1985) paper provided a detailed discussion of these events.

Series:

$y_t = 100 \cdot \ln(\text{GNP82}_t/\text{GNP82}_{t-1})$
$t = 1952\text{:III}-1984\text{:IV}$

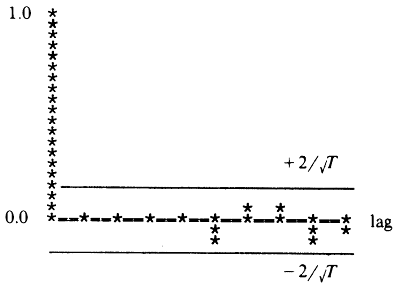
Sample Autocorrelogram:



Sample Regression Coefficients for AR(4) (Standard Errors in Parentheses):

$$y_t = \underset{(.129)}{.555} + \underset{(.089)}{.312 y_{t-1}} + \underset{(.093)}{.122 y_{t-2}} - \underset{(.092)}{.116 y_{t-3}} - \underset{(.089)}{.081 y_{t-4}} + u_t, \qquad \hat{\sigma}_u = 0.99$$

Sample Autocorrelogram of Residuals from AR(4) Regression:



Series

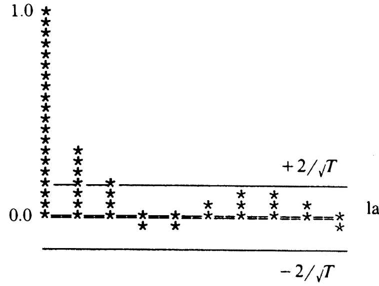
$y_t = -.3577 + 1.522 s_t + z_t$

$z_t = .014 z_{t-1} - 0.58 z_{t-2} - .247 z_{t-3} - .213 z_{t-4} + \varepsilon_t, \qquad \varepsilon_t \sim N[0, (.769)^2]$

$P[S_t = 1 | S_{t-1} = 1] = .9049$

$P[S_t = 0 | S_{t-1} = 0] = .7550$

Population Autocorrelogram:



Expected Value of Sample Regression Coefficients for AR(4):

$$y_t = .589 + .293 y_{t-1} + .069 y_{t-2} - .104 y_{t-3} - .042 y_{t-4} + u_t, \qquad \sigma_u = 0.98$$

Expected Value of Sample Autocorrelogram of Residuals from AR(4) Regression

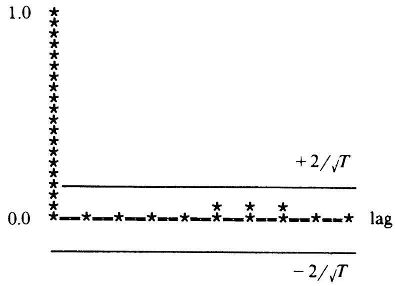


FIGURE 2.—Comparison of actual GNP data with predictions of Markov model of trend.

7. COMPARING LINEAR AND NONLINER MODELS OF GNP GROWTH

The Markov model offers a nonlinear alternative to linear representations such as the Box-Jenkins ARIMA specification (used by Beveridge and Nelson (1981), and Campbell and Mankiw (1987a, b)) or the unobserved components (UC) models of Harvey and Todd (1983), Watson (1986), and Clark (1987). One might well ask why, if the Markov model were the true data-generating process, do parsimoniously parameterized linear models seem to have fit the data so well?

Panel A of Figure 2 reports the sample autocorrelogram of actual postwar changes in the log of quarterly real GNP. Indeed this looks much like that predicted for low-order ARIMA processes.[10] An AR(4) model fit to the growth rate of real GNP exhibits only the most modest autocorrelation of residuals (Figure 2, panel A).

What would these same diagnostics be expected to reveal if the data were in fact generated by the Markov model? The Markov model posits that $y_t = \alpha_1 s_t + \alpha_0 + z_t$ with z_t a zero-mean Gaussian AR(r) process and $E(S_t) = \pi$. From the independence of S_t and z_t we know

$$E[y_t - Ey_t][y_{t-j} - Ey_{t-j}] = E[z_t z_{t-j}] + \alpha_1^2 E[S_t - \pi][S_{t-j} - \pi].$$

The first term is the jth autocovariance from a standard AR(r) process, and can be calculated using well-known formulas. Using (2.5) and the fact that $\mathrm{Var}(S_0) = \pi(1-\pi)$, we can evaluate the second term from

$$E[S_t - \pi][S_{t-j} - \pi] = \lambda^j \pi(1-\pi)$$

where as before $\lambda \equiv (-1 + p + q)$ and $\pi \equiv (1-q)/(1-\lambda)$. Thus the theoretical autocorrelogram of data generated by a Markov model is known. Panel B of Figure 2 plots this function for the MLE parameter values of Table I. It would clearly be extremely difficult to distinguish the Markov model from a simple linear alternative on the basis of the observed autocorrelations in a sample the size of postwar quarterly data.

Figure 2 also reports some Monte Carlo results. For each of 1000 samples of size $T = 130$ generated by the Markov model, an AR(4) specification was fit by OLS. The average regression coefficient vector across these samples (panel B) is very close to that for actual postwar data (panel A). The average sample autocorrelogram of the residuals again would provide negligible evidence against the AR(4) specification, even though we know that the true model used to simulate the data in panel B was the nonlinear Markov process and not an AR(4). I conclude that the Markov model satisfies the "encompassing" criterion of Hendry and Richard (1982)—the apparent success (on the basis of Box-Jenkins diagnostics) of simple ARIMA representations is precisely what one would predict if the Markov model were the true data-generating process.

There are, however, several predictions of the Markov model that are inconsistent with an ARIMA or linear UC specification. The Markov model asserts that forecasts of the log of GNP that are restricted to linear functions of lagged values will be suboptimal; additional useful information is alleged to be contained in the nonlinear function $P[S_{t-1} = 1 | y_{t-1}, y_{t-2}, \ldots]$ which summarizes the inference drawn the previous period about the unobserved state variable S_{t-1}. The intuition for the sign and magnitude of the predicted effect is as follows. If the Markov model were true and we knew that the economy was in the expansion phase of the cycle last period ($S_{t-1} = 1$), we would forecast

$$E[(\alpha_0 + \alpha_1 s_t) | S_{t-1} = 1] = \alpha_0 + \alpha_1 p$$

[10] For example, Watson (1986) settled on an ARIMA(1,1,0) specification for the log of GNP, Campbell and Mankiw (1987b) preferred a (2,1,2), and Clark (1987) selected (0,1,2).

whereas when the economy was in recession last period,

$$E[(\alpha_0 + \alpha_1 S_t)|S_{t-1} = 0] = \alpha_0 + \alpha_1(1 - q).$$

The difference in the forecast growth rate of the economy knowing that the economy was in expansion rather than recession last period would thus be on the order of $\alpha_1(p - 1 + q)$, or about 1% faster GNP growth forecast when the economy was in expansion last period.[11] The Markov model therefore predicts that the lagged output of the basic filter,

$$X_{t-1} \equiv P[S_{t-1} = 1 | y_{t-1}, y_{t-2}, \ldots, y_{-r+1}]$$

should enter statistically significantly with a positive coefficient when added to the AR(4) representation for GNP growth. The ARIMA or UC specifications predict that it should have coefficient zero. When one performs this regression on actual postwar GNP data, one finds (standard errors in parentheses)

$$y_t = -\underset{(.294)}{.199} - \underset{(.1609)}{.0721}\, y_{t-1} - \underset{(.10088)}{.00639}\, y_{t-2} - \underset{(.0927)}{.1815}\, y_{t-3}$$

$$- \underset{(.0917)}{.1639}\, y_{t-4} + \underset{(.590)}{1.670}\, X_{t-1} + u_t.$$

The t statistic associated with the null hypothesis that GNP growth rates were truly generated by an AR(4) model is 2.83, though X_{t-1} being a generated regressor, it is unclear what distribution theory is appropriate for interpreting this statistic.[12] The change in forecast is on the order of 1% of GNP.

Another prediction of the Markov model that is inconsistent with an ARIMA or UC specification concerns the heteroskedasticity of the residuals. The intuition is as follows. If the data were truly generated by the Markov model and we knew that the economy was in expansion last period ($S_{t-1} = 1$) along with knowing past values for ε_{t-j}, then the expected squared error in forecasting log GNP this period would be given by

$$E[\varepsilon_t^2] + E\{[(\alpha_1 S_t + \alpha_0) - (\alpha_1 p + \alpha_0)]^2 | S_{t-1} = 1\} = \sigma_\varepsilon^2 + \alpha_1^2 p(1 - p).$$

[11] This discussion (which rederives eq. (2.5) from first principles) is intended purely as an aid to the intuition. The formula in the text does not literally give the expected value of the coefficient in the regression that follows. One can of course arrive at the precise effect expected by adding s_{t-1} to the AR(4) regression of the Monte Carlo simulations described earlier. Its expected coefficient turns out to be 1.08.

[12] One might think it more natural to test the AR(4) specification against the Markov alternative as a conventional nested hypothesis. When $\alpha_1 = 0$, the growth rates in states 0 and 1 are the same. Thus, an AR(4) model for first-differences of the data obtains as a special case of the Markov specification, and one might think of using a likelihood ratio, Wald, or Lagrange multiplier test. Unfortunately, the usual regularity conditions for establishing asymptotic properties of these tests fail to apply here. Under the null hypothesis that $\alpha_1 = 0$, the parameters p and q are unidentified. When p, q, and α_1 are all treated as separate parameters, the information matrix is singular under the null hypothesis and the MLE's $\hat{p}$ and $\hat{q}$ cannot be regarded as consistent estimates of any population values. Furthermore, the derivative of the log likelihood with respect to α_1 is also zero at the constrained MLE. Davies (1977), Watson and Engle (1985), and Lee and Chesher (1986) have discussions of how one might try to construct asymptotic test statistics that are robust to these issues.

By contrast, if we knew that the economy was in recession last period,

$$E\left[\varepsilon_t^2\right] + E\left\{\left[(\alpha_1 S_t + \alpha_0) - (\alpha_1(1-q) + \alpha_0)\right]^2 | S_{t-1} = 0\right\}$$
$$= \sigma_\varepsilon^2 + \alpha_1^2 q(1-q).$$

Since $p > q > 1/2$, the model therefore predicts that an AR(4) forecast will have a smaller variance when the economy was in expansion last period than when the economy was in recession last period, the expected difference being on the order of[13]

$$\alpha_1^2\left[p(1-p) - q(1-q)\right] = -.229.$$

Thus, the Markov model predicts that in a regression of the square of the AR(4) residuals on a constant and the lagged filter output, the latter should enter statistically significantly and with a negative sign. The ARIMA or UC models predict homoskedastic errors and a coefficient of zero. In actual postwar GNP data one finds (standard errors in parentheses)

$$\hat{u}_t^2 = \underset{(.299)}{1.570} - \underset{(.369)}{.813}\, X_{t-1} + e_t, \qquad R^2 = .03654,$$

where $\hat{u}_t$ is the estimated residual from the AR(4) regression in panel B of Figure 2. The Breusch-Pagan (1979) test of the null hypothesis of homoskedastic errors is $1/\{2[(\hat{\sigma}_u^2)^2]\}$ times the explained sum of squares from this regression, which comes out to 5.17. Engle (1982, p. 1000) proposes calculating $TR^2 = 4.75$. Again abstracting from the generated regressor problem, both statistics should be $\chi^2(1)$ (whose 5% critical value is 3.84) under the null hypothesis that the data were generated by an AR(4) model with Gaussian homoskedastic errors. The data thus reveal evidence of the kind of conditional heteroskedasticity predicted by the Markov model and inconsistent with the ARIMA or UC specifications. Again the heteroskedasticity is economically large; (the squared residuals from an AR(4) are twice as large on average when the preceding period's inference about s_{t-1} pointed confidently to a recession).

8. ON THE CONSEQUENCES OF BUSINESS CYCLES FOR THE LONG RUN LEVEL OF OUTPUT

Much effort has recently been devoted to measuring the effect of an unanticipated increase in GNP on the optimal forecast of the level of GNP at an arbitrarily long time horizon. This question holds interest for two reasons. The first concerns the nature of the business cycle and its persistence; the second pertains to the response of consumers and firms to changing business conditions. I discuss the implications of my Markov parameterization for each of these issues in turn.

[13]Again, this discussion is meant primarily to highlight the intuition and not to derive the precise magnitude expected; the innovation of the AR(4) model is not simply $\varepsilon_t + (s_t - E_{t-1}S_t)$. From Monte Carlo simulations on data truly generated by the Markov model, the expected coefficient on s_{t-1} in the Breusch-Pagan regression that follows turns out to be $-.345$.

TABLE III

PREVIOUS ESTIMATES OF THE EFFECT OF AN UNANTICIPATED 1% INCREASE IN REAL GNP ON THE FUTURE LEVEL OF GNP AT AN ARBITRARILY LONG TIME HORIZON

BASIC WOLD REPRESENTATION: $(1-L)\tilde{y}_t = \mu + \psi(L)u_t$		$\psi(1)$
ARIMA$(p,1,q)$ *MODELS*: $\psi(L) = [1+\theta_1 L + \cdots + \theta_q L^q]/[1-\phi_1 L - \cdots - \phi_p L^p]$		
Watson (1986)	ARIMA(1,1,0)	1.68%
Clark (1987)	ARIMA(0,1,2)	1.62%
Campbell and Mankiw (1987b)	ARIMA(2,1,2)	1.49%
LINEAR UNOBSERVED COMPONENTS MODELS: $\psi(L)u_t = e_t^T + (1-L)\kappa(L)e_t^c$		
Watson (1986)		0.57%
Clark (1987)		0.64%
BIVARIATE MODEL: (univariate representation implied by bivariate process for GNP growth and level of unemployment)		
Evans (1987)	ARIMA(6,1,3)	0.55%
COCHRANE'S NONPARAMETRIC ESTIMATE:		
Campbell and Mankiw (1987a)		0.80% to 1.27%

8.1. *On the Nature and Persistence of the Business Cycle*

Nelson and Plosser (1982) and Campbell and Mankiw (1987a, b) were interested in the extent to which recessions represent temporary deviations from potential output with the shortfall largely made up during the subsequent recovery. Earlier approaches to this question were ultimately based on the standard linear representation for a nonstationary series $\tilde{y}_t$:

$$(1-L)\tilde{y}_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j} = \mu + \psi(L)u_t.$$

The permanent effect on the level of the series of a current innovation u_t is given by

$$\lim_{j\to\infty} \frac{\partial E_t \tilde{y}_{t+j}}{\partial u_t} = \sum_{j=0}^{\infty} \psi_j = \psi(1).$$

Previous researchers sought a finite-sample approximation to $\psi(L)$ based on Box-Jenkins methods, linear unobserved components models, bivariate models, and nonparametric tests. A sampling of estimates based on these techniques is provided in Table III.[14]

By contrast, the Markov model is fundamentally nonlinear and provides an alternative perspective on the basic question about business cycles posed by these researchers. We can write this model in the form

$$(1-L)\tilde{y}_t = (\alpha_0 + \alpha_1 s_t) + [\phi(L)]^{-1}\varepsilon_t.$$

Notice that the two fundamental sources of randomness, S_t and ε_t, are allowed to

[14] See also Cochrane (1987, 1988), Campbell and Deaton (1987), and Gagnon (1988). For comparison, the AR(4) model fit to GNP growth in panel 1 of Figure 2 implies $\psi(1) = 1.31$.

have very different implications for the future path followed by $\tilde{y}_t$. The earlier discussion argued that we could associate S_t with the business cycle directly, and ε_t with other factors contributing to changes in output. The permanent effect of the non-business-cycle component ε_t is given by

$$\lim_{j\to\infty} \frac{\partial E_t \tilde{y}_{t+j}}{\partial \varepsilon_t} = \frac{1}{\phi(1)} = \frac{1}{1 - .014 + .058 + .247 + .213} = 0.66.$$

On the other hand, if at date t the economy is in a recession ($S_t = 0$) rather than the growth state ($S_t = 1$), the consequences for the long-run future level of (100 times the log of) real GNP is given by equation (3.4):[15]

$$(5.1)\qquad \lim_{j\to\infty} \left\{ E_t\left[\tilde{y}_{t+j} | S_t = 1\right] - E_t\left[\tilde{y}_{t+j} | S_t = 0\right] \right\}$$
$$= \frac{\alpha_1(-1 + p + q)}{(2 - p - q)}$$
$$= \frac{1.522(-1 + .9049 + .7550)}{(2 - .9049 - .7550)} = 2.953$$

or about a 3% drop in GNP.

We can gauge the importance of Jensen's inequality for such calculations by using equation (3.16), which, in contrast to (5.1), forecasts the level rather than the log of GNP. Notice that for the MLE's in Table I, the term $\hat{\alpha}_1$ in equation (3.16) is estimated to be $\exp(1.522/100) = 1.01534$. The eigenvalues are $\mu_1 = 1.01138$ and $\mu_2 = 0.66264$. Thus from (3.16),

$$\lim_{j\to\infty} \left\{ E_t\left[\exp\left(\tilde{y}_{t+j}/100\right) | S_t = 1, z_t\right] \div E_t\left[\exp\left(\tilde{y}_{t+j}/100\right) | S_t = 0, z_t\right] \right\}$$
$$= \frac{1.01138 - (-1 + .9049 + .7550)}{1.01138 - (1.01534)(-1 + .9049 + .7550)} = 1.0297$$

virtually the identical 3% change predicted in eq. (5.1).

8.2. *Implications for the Permanent Income Hypothesis*

A conceptually separate reason for interest in the magnitudes in Table III arises from a desire to understand the spending habits of consumers. Here Deaton (1986) and Campbell and Deaton (1987) raise the issue as to whether an unanticipated 1% increase in income rationally signals a greater than 1% increase in permanent income. The magnitudes in Table III are then used to evaluate theories of consumption behavior as distinct from theories of the business cycle per se. Watson (1986) showed that different finite-parameter approximations to a

[15] This calculation holds the current level of GNP constant, and calculates only the "signalling" consequences of the recession for future GNP. If instead one wanted a dynamic multiplier (the future *and present* consequences of a shift from $S_t = 1$ to $S_t = 0$ with the history of ε's and all past s_{t-j} constant), one should add α_1 (or 1.522%) to the values reported in the text.

given process can yield strikingly different answers to this question. In this spirit I examine $\psi(1)$ for the linear Wold representation for my Markov process,

$$(8.1) \quad \begin{aligned} y_t &= \alpha_0 + \alpha_1 S_t + [\phi(L)]^{-1}\varepsilon_t \\ &= \mu + \psi(L)e_t. \end{aligned}$$

Note from (2.3), (2.13), and (8.1) that y_t has the spectrum

$$(8.2) \quad \begin{aligned} f(\omega) &= \frac{\sigma_\varepsilon^2}{(1-\phi_1 e^{i\omega} - \cdots - \phi_r e^{i\omega r})(1-\phi_1 e^{-i\omega} - \cdots - \phi_r e^{-i\omega r})} \\ &\quad + \frac{\alpha_1^2[p(1-p)\pi + q(1-q)(1-\pi)]}{(1-\lambda e^{i\omega})(1-\lambda e^{-i\omega})} \\ &= \sigma_e^2 \psi(e^{i\omega})\psi(e^{-i\omega}) \end{aligned}$$

where our task is to calculate $\psi(1)$. From (8.2) we see

$$\begin{aligned} f(0) &= \frac{\sigma_\varepsilon^2}{(1-\phi_1 - \cdots - \phi_r)^2} + \frac{\alpha_1^2[p(1-p)\pi + q(1-q)(1-\pi)]}{(1-\lambda)^2} \\ &= \sigma_e^2 \cdot [\psi(1)]^2. \end{aligned}$$

Using the maximum likelihood estimates in Table I, we calculate

$$(8.3) \quad \sigma_e^2 \cdot [\psi(1)]^2 = .261 + 2.277 = 2.538.$$

We further know (e.g., Anderson (1971, p. 422))

$$\sigma_e^2 = \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\omega)\, d\omega\right]$$

which one calculates to be .9703 by numerical integration of (8.2). Thus

$$\psi(1) = \left[\frac{\sigma_e^2 \cdot [\psi(1)]^2}{\sigma_e^2}\right]^{1/2} = 1.62.$$

This estimate is completely dominated by the contribution of the business cycle variable (see the second term in the sum on the right-hand side of (8.3)).

It is also straightforward to calculate the effect a recession would have on permanent income if consumers knew with certainty that a recession had started, that is, calculate the effect of a recession on the cumulative discounted value of future output flows. From equation (3.17), the ratio of the discounted value of the trend term when $\pi_0 = 1$ to the value when $\pi_0 = 0$ is given by[16]

$$\frac{1 - (-1+p+q)\beta \cdot \exp(\alpha_0/100)}{1 - (-1+p+q)\beta \cdot \exp[(\alpha_0+\alpha_1)/100]}.$$

[16]Recall that in the case of a Markov trend in logs, the stochastic specification is multiplicative, not additive ($\hat{y}_t = \hat{n}_t \hat{z}_t$) and so use of this formula is only strictly valid for $E_t \hat{z}_{t+j}$ constant. It does seem to offer a useful benchmark, however, for summarizing a key feature of these empirical estimates. See also the preceding footnote.

Using $\beta = 0.99$ for the quarterly real discount factor, this expression comes out to 1.029 for the empirical estimates in Table I; that is, the certain knowledge that the economy has gone into a recession is associated with a 3% drop in permanent income.

9. CONCLUSIONS

This paper explored the possibility that growth rates of real GNP are subject to autocorrelated discrete shifts. Empirical estimation suggested that the business cycle is better characterized by a recurrent pattern of such shifts between a recessionary state and a growth state rather than by positive coefficients at low lags in an autoregressive model. Indeed, statistical estimates of the economy's growth state cohere remarkably well with NBER dating of postwar recessions, and might be used as an alternative objective method for assigning business cycle dates. A move from expansion into recession is associated with a 3% decrease in the present value of future real GNP and similarly portends a 3% drop in the long-run forecast level of GNP.

Department of Economics, Rouss Hall, University of Virginia, Charlottesville, VA 22901, U.S.A.

Manuscript received November, 1986; final revision received June, 1988.

REFERENCES

ANDERSON, T. W. (1971): *The Statistical Analysis of Time Series*. New York: John Wiley and Sons, Inc.

AOKI, MASANAO (1967): *Optimization of Stochastic Systems: Topics in Discrete-Time Systems*. New York: Academic Press.

BEVERIDGE, STEPHEN, AND CHARLES R. NELSON (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151–174.

BOX, G. E. P., AND GWILYM M. JENKINS (1976): *Time Series Analysis: Forecasting and Control*, Revised Edition. San Francisco: Holden-Day.

BREUSCH, T. S., AND A. R. PAGAN (1979): "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*, 47, 1287–1294.

BROCK, W. A., AND CHERA L. SAYERS (1988): "Is the Business Cycle Characterized by Deterministic Chaos?" *Journal of Monetary Economics*, 22, 71–80.

CAMPBELL, JOHN Y., AND ANGUS DEATON (1987): "Is Consumption Too Smooth?" NBER Working Paper No. 2134.

CAMPBELL, JOHN Y., AND N. GREGORY MANKIW (1987a): "Permanent and Transitory Components in Macroeconomic Fluctuations," *American Economic Review Papers and Proceedings*, 77, 111–117.

——— (1987b): "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 102, 857–880.

CHIANG, CHIN LONG (1980): *An Introduction to Stochastic Processes and Their Applications*. New York: Krieger.

CLARK, PETER K. (1987): "The Cylical Component of U.S. Economic Activity," *Quarterly Journal of Economics*, 102, 797–814.

COCHRANE, JOHN H. (1987): "Spectral Density Estimates of Unit Roots," Working Paper, University of Chicago.

——— (1988): "How Big is the Random Walk in GNP?" *Journal of Political Economy*, 96, 893–920.

COSSLETT, STEPHEN R., AND LUNG-FEI LEE (1985): "Serial Correlation in Discrete Variable Models," *Journal of Econometrics*, 27, 79–97.

DAVIES, R. B. (1977): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247–254.

DEATON, ANGUS S. (1987): "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in *Advances in Econometrics, Fifth World Congress*, Vol. II., ed. by T. F. Bewley. New York: Cambridge University Press, pp. 121–148.

DIEBOLD, FRANCIS X., AND GLENN D. RUDEBUSCH (1987): "Scoring the Leading Indicators," Federal Reserve Board, Special Studies Paper No. 206.

ENGLE, ROBERT F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

ENGLE, ROBERT F., AND C. W. J. GRANGER (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

ENGLE, ROBERT F., DAVID LILIEN, AND RUSSELL P. ROBINS (1987): "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 55, 391–407.

EVANS, GEORGE W. (1987): "Output and Unemployment Dynamics in the United States: 1950–1985," Working Paper, Stanford University.

GAGNON, JOSEPH E. (1988): "Short-Run Models and Long-Run Forecasts: A Note of the Permanence of Output Fluctuations," *Quarterly Journal of Economics*, 103, 415–424.

GALLANT, A. RONALD, AND GEORGE TAUCHEN (1987): "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," Mimeographed, North Carolina State University.

GOLDFELD, STEPHEN M., AND RICHARD E. QUANDT (1973): "A Markov Model for Switching Regressions," *Journal of Econometrics*, 1, 3–16.

GRANGER, C. W. J. (1983): "Forecasting White Noise," *Applied Time Series Analysis of Economic Data*, in *Proceedings of the Conference on Applied Time Series Analysis of Economic Data*, Oct. 13–15, 1981, Arlington, VA, ed. by Arnold Zellner. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, pp. 308–314.

HAMILTON, JAMES D. (1985): "Historical Causes of Postwar Oil Shocks and Recessions," *Energy Journal*, 6, 97–116.

——— (1988): "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control*, 12, 385–423.

HARVEY, A. C. (1985): "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3, 216–227.

HARVEY, A. C., AND P. H. J. TODD (1983): "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study," *Journal of Business and Economic Statistics*, 1, 299–307.

HENDRY, DAVID F., AND JEAN-FRANCOIS RICHARD (1982): "On the Formulation of Empirical Models in Dynamic Econometrics," *Journal of Econometrics*, 20, 3–33.

HINICH, MELVIN J., AND DOUGLAS M. PATTERSON (1985): "Evidence of Nonlinearity in Daily Stock Returns," *Journal of Business and Economic Statistics*, 3, 69–77.

KING, ROBERT, CHARLES PLOSSER, JAMES STOCK, AND MARK WATSON (1987): "Stochastic Trends and Economic Fluctuations," NBER Working Paper No. 2229.

LEE, LUNG-FEI, AND ANDREW CHESHER (1986): "Specification Testing when the Score Statistics Are Identically Zero," *Journal of Econometrics*, 31, 121–149.

LIPTSER, R. M., AND A. N. SHIRYAYEV (1977): *Statistics of Random Processes, Volume I: General Theory*. New York: Springer-Verlag.

NEFTCI, SALIH N. (1982): "Optimal Prediction of Cyclical Downturns," *Journal of Economic Dynamics and Control*, 4, 225–241.

——— (1984): "Are Economic Time Series Asymmetric over the Business Cycle?" *Journal of Political Economy*, 92, 307–328.

NELSON, CHARLES R., AND CHARLES I. PLOSSER (1982): "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–162.

QUAH, DANNY (1986): "What Do We Learn from Unit Roots in Macro Economic Time Series?" Working Paper, MIT.

SCLOVE, STANLEY L. (1983): "Time-Series Segmentation: A Model and a Method," *Information Sciences*, 29, 7–25.

SICHEL, DANIEL E. (1987): "Business Cycle Asymmetry: A Deeper Look," Mimeographed, Princeton University.

STOCK, JAMES H. (1987): "Measuring Business Cycle Time," *Journal of Political Economy*, 95, 1240–1261.

TONG, HOWELL (1983): *Threshold Models in Non-linear Time Series Analysis*. New York: Springer-Verlag.

WATSON, MARK W. (1986): "Univariate Detrending Methods with Stochastic Trends," *Journal of Monetary Economics*, 18, 49–75.

WATSON, MARK W., AND ROBERT F. ENGLE (1985): "Testing for Regression Coefficient Stability with a Stationary AR(1) Alternative," *Review of Economics and Statistics*, 67, 341–346.

WECKER, WILLIAM E. (1979): "Predicting the Turning Points of a Time Series," *Journal of Business*, 52, 35–50.