Journal of Business & Economic Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ubes20

Business-Cycle Analysis With a Markov-Switching Model

Thomas H. Goodwin [a]
[a] Program in Economics, Claremont Graduate School, Claremont, CA, 91711
Published online: 02 Jul 2012.

Journal of Business & Economic Statistics, July 1993, Vol. 11, No. 3

Business-Cycle Analysis With a Markov-Switching Model

Thomas H. Goodwin
Program in Economics, Claremont Graduate School, Claremont, CA 91711

This article explores the Hamilton Markov-switching model through an analysis of the business cycles of eight developed market economies. Forecasting and specification tests suggest only marginal improvements over linear autoregressive models. Yet filtered and smoothed conditional probabilities indicate turning points in business cycles that closely correlate with turning points from traditional methods. Tests regarding the asymmetry of business cycles reject the null of symmetry for most countries.

KEY WORDS: Dating peaks and troughs; Markov in the mean; Mixture normal; Nonlinear time series models; Regime shifts; Switching regressions.

Characterizing the nature of business cycles and long-term growth in aggregate output has long been an objective of empirical macroeconomics. In the last dozen years, several new techniques have been employed that have produced new insights. One approach, used by Beveridge and Nelson (1981), Nelson and Plosser (1982), and Campbell and Mankiw (1987), emphasizes autoregressive integrated moving average models. Another discussed by Harvey (1985), Watson (1986), and Clark (1987), models real gross national product (GNP) as the sum of unobserved components using the Kalman filter. Yet a third approach, that of King, Plosser, Stock, and Watson (1991), exploits cointegrating relations across several macro time series.

Although all of the preceding research has produced contributions to our understanding, they also share a possible shortcoming—the assumption that the growth rate of real GNP is a linear stationary process. Linear models are incompatible with the asymmetry between expansions and contractions that has been documented by, among others, Neftci (1984), Stock (1987), Diebold and Rudebusch (1990), and Sichel (1993). Moreover, evidence of more general departures from linearity in many macroeconomic and financial time series has begun to mount in a wide range of works such as Engle, Lilien, and Robins's (1987) autoregressive conditional heteroscedastic in the mean (ARCH-M) process, Brock and Sayers's (1988) chaos model, and Hinich and Patterson's (1985) bispectral analysis.

Recently, Hamilton (1989) proposed a nonlinear alternative to the class of linear stationary models that holds promise as a new method for testing hypotheses about and dating business cycles. This article is an exploration of this new technique through an analysis of real GNP for eight developed countries in the postwar era.

The article is organized as follows. Section 1 sets out the basic Hamilton model, presents estimates for eight Organization for Economic Cooperation and Development (OECD) countries, and conducts a specification analysis. Section 2 presents estimated dates for turning points in the eight countries and compares them to several other dating methods. Section 3 uses the model estimates to test the asymmetry hypothesis about business cycles. Section 4 concludes the article.

1. ESTIMATION AND SPECIFICATION TESTING OF THE HAMILTON MODEL

The Hamilton (1989) model specifies that real GNP growth follows an (autoregressive) AR(4) process. Nonlinearity of the model arises because the process is subject to discrete shifts in the mean, between *high-growth* and *low-growth* states. These discrete shifts have their own dynamics, specified as a two-state first-order Markov process. The model is written

$$y_t - \mu_{s_t} = \phi_1(y_{t-1} - \mu_{s_{t-1}}) + \phi_2(y_{t-2} - \mu_{s_{t-2}}) + \phi_3(y_{t-3} - \mu_{s_{t-3}}) + \phi_4(y_{t-4} - \mu_{s_{t-4}}) + \sigma\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \tag{1}$$

$$\mu_{s_t} = \alpha_0 + \alpha_1 s_t, \tag{2}$$

$s_t = 1$ if high growth state, 0 otherwise,

$$P[S_t = 1|S_{t-1} = 1] = p, \tag{3}$$

$$P[S_t = 0|S_{t-1} = 1] = 1 - p,$$

$$P[S_t = 0|S_{t-1} = 0] = q,$$

and

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

The model is a nonlinear combination of discrete and continuous dynamics with nine parameters: (α_1, α_0, p, q, σ, ϕ_1, ϕ_2, ϕ_3, ϕ_4). Since $y_t = 100 \cdot \ln(\text{RealGNP}_t/\text{RealGNP}_{t-1})$, α_0 is the mean quarterly percentage rate of growth in the low-growth state, and $\alpha_0 + \alpha_1$ is the mean quarterly percentage rate of growth in the high-growth state. An attractive feature of the model is that no prior information regarding the dates of the two growth periods or the size of the two growth rates is required. In particular, note that the low-growth rate need not be negative. Derivation of the sample conditional log-likelihood $\Sigma_{t=1}^{T} \ln f(y_t|y_{t-1}, y_{t-2}, \ldots, y_{-3})$ was detailed by Hamilton (1989) and will not be repeated here. Henceforth, the Hamilton model will be referred to as the MS(4) model, for Markov switching in the mean with an AR(4) component.

Data on real GNP were gathered from the International Financial Statistics tape of the International Monetary Fund for the United States, the United Kingdom, Germany, Japan, Canada, Switzerland, Italy, and France. Maximum likelihood estimation of the nine parameters for each of the eight countries was performed using the Optmum module of GAUSS 2.2 with a combination of the Davidon–Fletcher–Powell and steepest ascent numerical algorithms. Constraints were placed on p and q to restrict them to the 0–1 interval, and σ was restricted to the positive half of the real line.

The likelihood function of the MS(4) model is typically ill behaved with numerous local optima. Starting values can have a profound impact on which local optimum is found. To improve the chances of finding the global optimum, a grid of 720 points in the space of α_1, p, and q is specified. The likelihood is optimized with respect to the other parameters at each fixed point of the (α_1, p, q) grid. The exact grid will be given later. The parameter values of the supremum of the 720 maximized likelihoods are then used as starting values for finding the optimum with respect to all nine parameters jointly. This approach has a high probability of producing a global maximum.

For four of the countries—the United Kingdom, Japan, Italy, and France—the global optimum resulted in one of the transition probabilities, either p or q, approaching 0. When either p or q are close to 0, the model is useless for dating turning points in the business cycle because only one state persists through most of the sample period. The problem is obviated for three of the countries by employing the "quasi-Bayesian" prior of Hamilton (1991). This prior is in the same spirit as the Theil–Goldberger "mixed" estimator in that pseudo observations are created with a prespecified location and precision. In all cases in which a prior is used here, it is equivalent to adding one additional observation from each state with zero mean and unit variance. For one country, Italy, estimation with this prior still resulted in a small value for p, although it did help in obtaining a quick convergence. The estimates are presented in Table 1.

Table 2 contains the results from a battery of specification and forecasting tests. Comparison is made with maximum likelihood estimates (not presented) of a linear autoregressive model of fourth order, an AR(4) without a switching mean:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \phi_3(y_{t-3} - \mu) + \phi_4(y_{t-4} - \mu) + \sigma\varepsilon_t, \qquad \varepsilon_t \sim N(0, 1). \quad (4)$$

The rationale for an AR(4) as a competing model is that virtually all linear stationary processes have an AR

Table 1. Maximum Likelihood Estimates of the Hamilton Markov-Switching Model for Eight Economies

	U.S.	U.K.	Germany	Japan	Canada	Switzerland	Italy	France
α_0	−.39	−.44	−.27	−.25	−1.06	−3.07	.75	−1.074
	(.39)	(.30)	(.35)	(.12)	(.58)	(.59)	(.17)	(.49)
α_1	1.45	1.25	1.32	1.88	2.19	3.85	4.91	2.026
	(.34)	(.31)	(.32)	(.86)	(.58)	(.58)	(.62)	(.47)
p	.94	.96	.91	.98	.99	.96	.0057	.97
	(.033)	(.031)	(.053)	(.034)	(.010)	(.023)	(.046)	(.02)
q	.73	.79	.71	.64	.74	.60	.97	.51
	(.15)	(.14)	(.17)	(.30)	(.22)	(.17)	(.015)	(.27)
σ	.77	1.28	1.15	1.17	.96	1.29	1.21	1.11
	(.060)	(.094)	(.10)	(.094)	(.063)	(.12)	(.08)	(.086)
ϕ_1	.076	−.34	−.30	−.12	.064	−.31	.34	−.49
	(.11)	(.10)	(.11)	(.11)	(.091)	(.12)	(.11)	(.11)
ϕ_2	.046	−.12	−.18	.29	.025	−.080	.13	.14
	(.11)	(.11)	(.13)	(.86)	(.087)	(.12)	(.11)	(.11)
ϕ_3	−.16	−.13	.065	.28	.13	.21	.01	.49
	(.095)	(.11)	(.14)	(.12)	(.093)	(.12)	(.11)	(.13)
ϕ_4	−.062	−.14	.30	.21	−.20	.30	−.17	.28
	(.092)	(.10)	(.13)	(.13)	(.091)	(.12)	(.10)	(.13)
log L	−49.01	−103.56	−88.78	−89.29	−65.10	−65.40	−82.31	−66.37
Q-B prior?	No	Yes	No	Yes	No	No	Yes	Yes
T	132	130	119	131	130	75	116	99
Period	57:2–90:1	57:2–89:3	60:2–89:4	57:2–89:4	57:2–89:3	67:2–85:4	61:1–89:4	65:2–89:4

NOTE: Standard errors are in parentheses.

Table 2. Specification Tests and Forecast Statistics

Models	U.S.	U.K.	Germany	Japan	Canada	Switzerland	Italy	France
				Hansen standardized LR test—H_0: AR(4) versus H_1: MS(4)				
	1.81	1.11	1.28	1.23	1.34	1.75	1.79	1.45
	(.56)	(.92)	(.80)	(.79)	(.79)	(.49)	(.46)	(.64)
				Nyblom stability test				
MS(4)	1.14	3.12*	2.19*	2.86*	1.37	2.30*	4.04*	4.06*
AR(4)	.62	.47	1.33*	1.23	1.40*	1.01	.82	1.62*
				Within-sample RMSE				
MS(4)	.91	1.35	1.32	1.23	1.05	1.70	1.51	1.31
AR(4)	.99	1.48	1.39	1.27	1.16	1.72	1.59	1.35
				Within-sample MAE				
MS(4)	.71	1.03	1.04	.85	.80	1.17	.89	.73
AR(4)	.79	1.14	1.08	.90	.88	1.27	.99	.74
				Within-sample Theil's U_Δ				
MS(4)	1.12	.88	.88	.94	1.02	.97	1.11	.81
AR(4)	1.22	.98	.94	.99	1.14	.95	1.10	.82
				Out-of-sample RMSE				
MS(4)	.46	.73	1.18	1.30	.53	.73	.25	.40
AR(4)	.35	.71	1.22	1.33	.44	.83	.38	.39
				Coefficient of $P(s_{t-1} = 1 \mid Y_{t-1})$ in an AR(4) regression				
	.93	1.71	1.05	.96	.57	1.07	.25	−4.88
				Coefficient of $P(s_{t-1} = 1 \mid Y_{t-1})$ in a regression with ε_t^2 as dependent variable				
	−1.25	−.17	−.49	−3.65	−4.70	−1.59	−.71	−14.83
				Auxiliary regression test of serial correlation up to fourth order				
MS(4)	.38	.124	.071	.21	.089	.23	.76	1.75
	(.83)	(.97)	(.99)	(.93)	(.99)	(.92)	(.56)	(.15)
AR(4)	1.173	2.44	.45	.58	3.37	.16	1.20	2.38
	(.33)	(.051)	(.76)	(.68)	(.012)	(.96)	(.31)	(.058)
				Auxiliary regression test of an ARCH process up to fourth order				
MS(4)	1.06	1.18	2.43	1.94	.46	2.14	1.45	7.32
	(.38)	(.32)	(.052)	(.11)	(.77)	(.087)	(.22)	(.0001)
AR(4)	.72	.34	3.60	4.44	.30	2.43	1.25	18.88
	(.58)	(.85)	(.009)	(.002)	(.88)	(.057)	(.29)	(.0001)

NOTE: p values in parentheses.
*Rejects null of parameter stability for a test of size 5%.

representation in which ε_t is approximately standard Gaussian white noise. Moreover, the AR(4) model is nested within the MS(4) model. The discussion of the tests proceeds in the order of presentation in Table 2.

An important question is whether the apparent good fit of the MS(4) model, as suggested by the large differences between α_0 and α_1 for the eight countries, can be explained by sampling error. The MS(4) model reduces to the AR(4) model under the restriction

$$H_0\colon \alpha_1 = 0. \tag{5}$$

It would be tempting to apply a conventional t test or likelihood ratio (LR) test for the significance of α_1. It is easy to see that such a test would reject (5) for all eight countries. These test statistics, however, do not have standard null distributions because, under the null hypothesis, the transition probabilities p and q are not identified. A score test based on the AR(4) estimates is not available either, since the scores with respect to α_1, p, and q, are identically 0 under the null. This leaves one with the choice of either conducting a Monte Carlo simulation as in the work of Lam (1990) or employing the approach of Hansen (1992). The latter is chosen here.

Hansen (1992) drew on empirical process theory to view the LR statistic as the sum of a limit function and an empirical process under the null. Under fairly general conditions, the empirical process is mean-zero Gaussian with a covariance function dependent on nuisance parameters. Because of this dependency, the LR must be standardized for analysis. The supremum of the standardized LR statistics calculated over the entire space of (α_1, p, q) is chosen for the test statistic.

The testing procedure consists of several steps. First, a large grid in the space of (α_1, p, q) is specified; for α_1, the range of [.3, 6] in steps of .3 is chosen, and for p and q the range of [.15, .90] in steps of .15. This partitions the space into 720 gridpoints. Next, the likelihood is maximized with respect to the "nuisance" parameters (α_0, ϕ_1, ϕ_2, ϕ_3, ϕ_4, σ) for each fixed gridpoint. At each optimum, an LR statistic is calculated under the null of AR(4) versus the alternative of MS(4), and standardized. Next, the supremum of the 720 standardized LR statistics is chosen for the test statistic. These values are presented in the first row of Table 2. The asymptotic distribution of the statistic is dependent on the data and parameters, so generic tabulation is not possible. The second row of Table 2 presents p values calculated from using 10,000 draws from the normal in conjunction with the empirical covariance functions.

The simulated p values of the standardized LR tests are uniform in asserting that the difference between the MS(4) and AR(4) models cannot be distinguished from sampling error, reversing the results of conventional tests. For four of the countries, this result was signaled by the tendency of one of the transition probabilities to approach 0 without imposing a quasi-Bayesian prior. Hansen (1992) pointed out that his testing procedure results in conservative p values—that is, upper bounds for the true p values. He conducted a study of the power of the procedure and found that the test has high power. But he only investigated power in the case of no AR component, so caution is warranted in drawing conclusions. Nonetheless, this is a striking result.

The next test in Table 2 is the parameter stability test of Nyblom (1989) based on the cumulative sums of the score function. The starred entries indicate a rejection of the null of stability at the 5% critical values that Nyblom tabulated. If the AR(4) models generally indicated instability while the MS(4) models indicated stability, it would be a powerful argument in favor of Markov switching. Such is not the case, however, because six of the MS(4) models reject, but only three of the AR(4) reject.

The next four sets of statistics in Table 2 are based on forecasts. Forecasts from an AR(4) model are straightforward and do not require comment. The within-sample forecasts for the MS(4) models are based on the one-step-ahead formula

$$\begin{aligned} E(y_{t+1}|\mathbf{Y}_t) &= E(\mu_{s_{t+1}}|\mathbf{Y}_t) + \sum_{i=1}^{4} \phi_i E[(y_{t-i+1} - \mu_{s_{t-i+1}})|\mathbf{Y}_t] \\ &= \mathbf{a}'_{t|t}\mathbf{P}\boldsymbol{\mu} + \sum_{i=1}^{4} \phi_i(y_{t-i+1} - \mathbf{a}'_{t-i+1|t}\boldsymbol{\mu}), \end{aligned} \quad (6)$$

where

$$\mathbf{a}_{\tau|t} = \begin{pmatrix} P(s_\tau = 1|\mathbf{Y}_t) \\ P(s_\tau = 0|\mathbf{Y}_t) \end{pmatrix},$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_0 \end{pmatrix} = \begin{pmatrix} \alpha_0 + \alpha_1 \\ \alpha_0 \end{pmatrix},$$

$$\mathbf{P} = \begin{pmatrix} p & (1-p) \\ (1-q) & q \end{pmatrix},$$

$\mathbf{Y}_t = (y_t, y_{t-1}, y_{t-2}, \ldots, y_{-3})$, $\boldsymbol{\mu}$ is the vector of means for the two states, $\mathbf{P}$ is the Markov-probability transition matrix, and $\mathbf{Y}_t$ represents all observable information at time t. The optimal one-step-ahead forecast of y will use both the autoregressive dynamics of the observed current and past y and the Markovian dynamics of the unobserved states. The conditional probabilities, $P(s_\tau = k|\mathbf{Y}_t)$, are constructed post-estimation using the nonlinear filtering method of Hamilton (1989, in press).

For each of the eight countries three prediction statistics for one-step-ahead forecasts are presented—root mean squared error (RMSE), mean absolute error (MAE), and Theil's U for first differences. In the standard modeling situation, one would expect that the MS(4) model would have smaller within-sample RMSE than an AR(4) by construction, since the MS(4) has more parameters and encompasses the AR(4). This is not a standard situation, however. Forecasts from the MS(4) model are a function of the conditional probabilities, which are not an explicit part of the likelihood function. Thus there is no guarantee regarding RMSE between the two models. Indeed, Phillips (1991) found that an AR(1) model had a lower within-sample RMSE than an MS(1) for the industrial production of three of four countries.

There are at least two versions of Theil's statistic in the literature; the one used here is

$$U_\Delta = \{(1/T)\Sigma[\Delta y_{t+1} - \Delta E(y_{t+1}|\mathbf{Y}_t)]^2\}^{1/2} \div \{(1/T)\Sigma\Delta y_{t+1}^2\}^{1/2}. \quad (7)$$

The U_Δ statistic picks up the ability of the model to anticipate turning points.

The within-sample forecast statistics, the RMSE and MAE, are uniformly smaller for the MS(4) model. The average reduction in RMSE is about 6%. The U_Δ statistics indicate that the MS(4) models anticipate turning points better for six of eight countries.

Although within-sample forecast statistics are useful, many would argue that out-of-sample forecasts are more telling. For each of eight countries, the last four quarters of GNP are retained and the models reestimated. One- through four-step-ahead forecasts are generated from the MS(4) models using the autoregressive-Markovian recursion

$$E(y_{t+m}|\mathbf{Y}_t) = \mathbf{a}'_{t|t}\mathbf{P}^m\boldsymbol{\mu} + \sum_{i=1}^{4} \phi_i[E(y_{t+m-i}|\mathbf{Y}_t) - \mathbf{a}'_{t|t}\mathbf{P}^{m-i}\boldsymbol{\mu}], \quad (8)$$

where $\mathbf{a}'_{t|t}\mathbf{P}^{m-i}\boldsymbol{\mu} = \mathbf{a}'_{t-i+1|t}\boldsymbol{\mu}$, if $m - i \leq 0$. The RMSE's of the out-of-sample forecasts divide evenly: Four countries have smaller RMSE's from the MS(4) model and four have smaller RMSE's from the AR(4) model.

Hamilton (1989, pp. 376–378) argued that, if the data were in fact generated by his model, then the lagged conditional probability of the high-growth state, if used as a regressor in a linear AR(4) model like (4), should have a positive and significant coefficient. Such regressions were run, and their estimated coefficients are presented in Table 2. All coefficients have the anticipated sign except France. Because the conditional probability is a generated regressor, consistent standard errors are not available unless bootstrapped. So "significance" is left as an open question.

Hamilton also pointed out that the variance of the Markov model is heteroscedastic. If an AR(4) model is estimated with the assumption of homoscedastic errors, then the lagged conditional probability of the high-growth state should have a negative coefficient in a simple regression with ε_t^2 from the AR(4) as dependent variable. Those coefficients are also presented in Table 2. All of the coefficients have the expected sign if the data were indeed generated by an MS(4) model.

Finally, the four bottom rows of Table 2 contain results from auxiliary regression tests for serial correlation and unaccounted-for (autoregressive conditional heteroscedasticity) ARCH/nonlinear effects. For these tests, the forecast errors must be normalized by the forecast standard deviations, since the errors are time-varying heteroscedastic in the MS(4) model. The one-step-ahead forecast error and variance is calculated by extending the formulas of Hamilton (1989) and Engle and Hamilton (1990). The normalized forecast errors, $\varepsilon_{t+1|t}/\sigma_{t+1|t}$, are regressed on a constant and four own-lags as a test for serial correlation. The squares of the normalized errors are also regressed on four own-lags of the squared errors as a test for an unspecified ARCH/nonlinear process. Table 2 presents F statistics and p values for the significance of the regressions. Kiviet (1986) showed that the F statistic has better small-sample properties than the more familiar TR^2. Normalized one-step-ahead forecast errors from the AR(4), $\varepsilon_{t+1|t}/\sigma$, are used to calculate analogous statistics for comparison.

None of the auxiliary regression statistics show evidence of serial correlation in the MS(4) models, but three of eight indicate serial correlation in the AR(4) model at the 10% level of significance. Three of eight MS(4) models show signs of an ARCH/nonlinear process in the errors at the 10% level, whereas four of the AR(4) models show such signs. In those cases in which the MS(4) model has significant ARCH/nonlinear errors, there still is some improvement over a linear model. Nonetheless, the rejections indicate that the MS(4) model does not capture all nonlinearities in the data.

Overall, this battery of tests suggests that the Hamilton specification does not strongly dominate linear representations of GNP growth. The most damaging evidence is the Hansen LR test, which does not come close to rejecting the null of AR(4) for any of the eight countries. On the other hand, several of the other tests indicate that the MS(4) has something to contribute in explanatory power. In Section 2, the most innovative feature of the Hamilton model is examined—the ability to pinpoint probable turning points of an unobserved process.

2. DATING TURNING POINTS WITH THE HAMILTON MODEL

There are two types of decision-rule systems for identifying business-cycle turning points. The most well known was developed by Mitchell and Burns (1938) and subsequent colleagues at the National Bureau of Economic Research (NBER). The methodology relies on indexes of coincident, leading, and lagging indicators and a series of *signals* or *checkpoints* to spot and confirm turning points. A fair amount of judgment is involved. A compilation of NBER turning points for the United States is published regularly by the Department of Commerce. Many researchers take these dates as definitive, but others dismiss them as "measurement without theory" (Koopmans 1965).

A second type of decision-rule system is based on a probability model. Examples are the Neftci (1982) sequential analysis method and the Stock and Watson (1991) single-index method. A drawback of these probability models is that they use prior information regarding turning-point dates, usually the NBER dates themselves, or some other ad hoc rule, so they are not truly independent of the NBER methods. The Hamilton Markov-switching model is a probability model with the important difference that no prior information on turning points is required. Conditional probabilities of expansionary and contractionary growth phases can be constructed post-estimation to suggest turning points. Indeed, the most striking result of Hamilton (1989) was the ability of the model to pinpoint turning points very close to the NBER dates without any prior restrictions.

Practitioners of the traditional methods of business-cycle dating have identified two types of business cycles. The *classical* cycle is an absolute decline and recovery in economic activity. This is the type that the NBER dates for the United States. Applying that standard to OECD countries, however, results in few classical cycles for many of these economies during the postwar period.

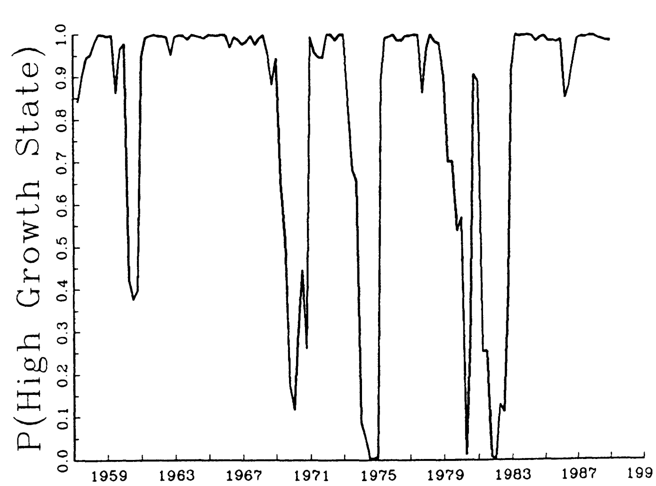

Figure 1. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for the United States.

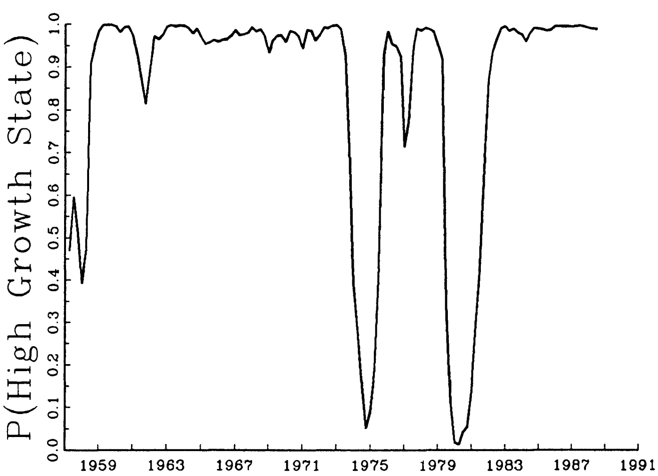

Figure 2. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for the United Kingdom.

Consequently, many OECD governments prefer to date "growth" cycles, positive and negative deviations around a trend rate of growth. The Hamilton model permits either classical cycles, $\alpha_0 < 0$, or growth cycles, $\alpha_0 > 0$. The estimates in Table 1 indicate classic turning points for all countries except in the troublesome case of Italy.

Which state an economy is in at any particular time remains unobservable in the Hamilton framework, as in reality. The probability of a particular state at a particular time can be calculated based on the estimated parameters and the path of the process through the use of filtered inferences. The precise steps were outlined by Hamilton (1989, in press). The filtering process involves forming the joint conditional probability $P[S_t, S_{t-1}, S_{t-2}, S_{t-3}, S_{t-4}, y_t|\mathbf{Y}_{t-1}]$. By integrating out the other states and dividing by the fitted likelihood, the conditional probability of a particular state, say state 1, can be obtained: $P[S_t = 1|\mathbf{Y}_t]$. The time path of conditional probabilities calculated this way is often volatile. Hamilton (1989) suggested using forward information on the y_t process to, in effect, smooth the path of probabilities. A *four-lag smoother* is employed here. This involves integrating out the forward states instead of the past states: $P[S_{t-4} = 1|\mathbf{Y}_t]$. Shifting the time subscripts forward by four produces $P[S_t = 1|\mathbf{Y}_{t+4}]$, the four-lag smoother.

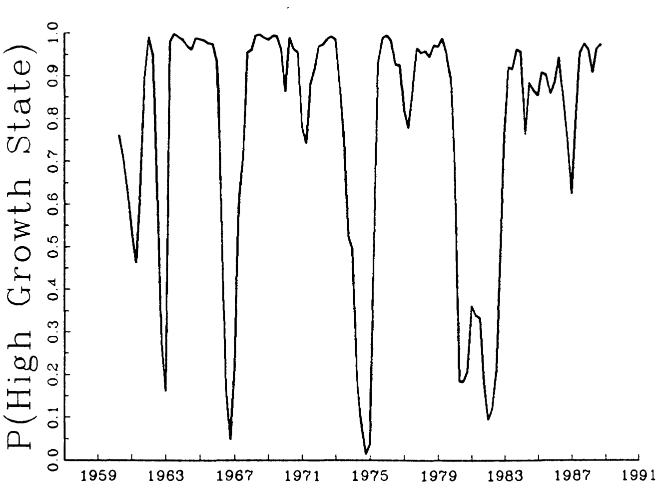

Figure 3. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for Germany.

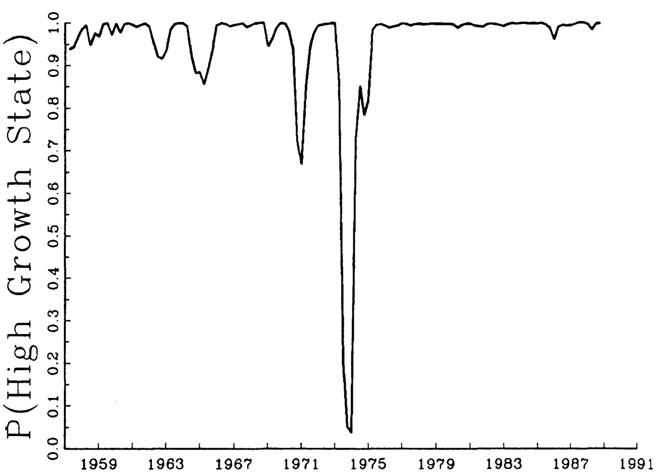

Figure 4. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for Japan.

Time paths of conditional probabilities of the high-growth state were calculated from the estimates of Table 1 for each of the eight economies using the smoothed and filtered inferences. Figures 1–8 plot these estimated probabilities. It is natural to view the high-growth state as an expansionary phase and the low-growth state as a recessionary phase.

As mentioned previously, Italy is a case in which the cyclical interpretation of high and low growth states fails spectacularly, and Figure 7 confirms that. There are three very large outliers that the likelihood identifies as the high-growth state—1966:1, 1970:1, and 1971:4. All else is relegated to the low-growth state. Various forms of data-smoothing techniques were tried to ren-

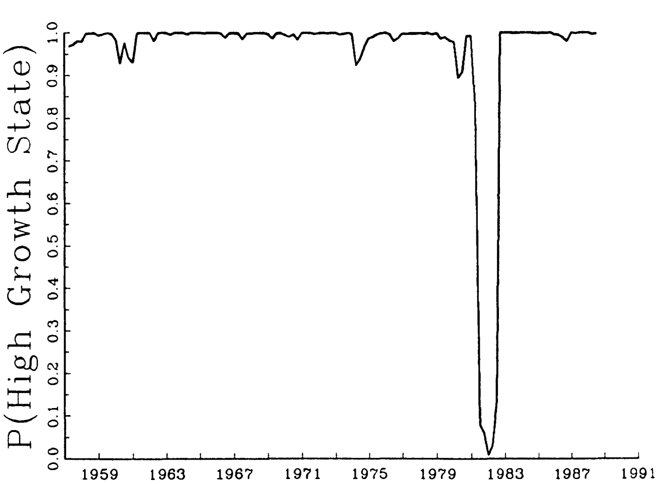

Figure 5. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for Canada.

Figure 6. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for Switzerland.

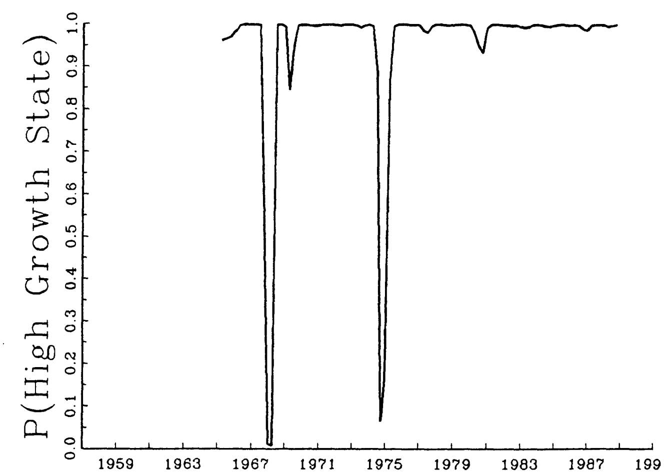

Figure 8. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for France.

der the outliers less influencial, such as year-over-year growth and weighted moving averages. The result was conditional probabilities that wandered around the range of .4 to .6. Apparently, smoothing the data smears the information content enough that the likelihood cannot distinguish between the two states very precisely. A heavier quasi-Bayesian prior was also imposed, to no avail. Since these estimates will be useless for dating business cycles, the Italian case will be excluded from the rest of the analysis.

Hamilton (1989) suggested using the rule that a quarter is part of a low-growth, or recessionary, period if $P[S_t = 1|\mathbf{Y}_{t+4}] < .5$; that is, the economy is most likely in the low-growth state. The first column of Table 3 contains dates of turning points for the United States from the MS(4) model. Using traditional terminology, a date τ is designated a *peak* if $P[S_\tau = 1|\mathbf{Y}_{\tau+4}] > .5$ and $P[S_{\tau+1} = 1|\mathbf{Y}_{\tau+5}] < .5$. Likewise, a date τ is designated a *trough* if $P[S_\tau|\mathbf{Y}_{\tau+4}] < .5$ and $P[S_{\tau+1}|\mathbf{Y}_{\tau+5}] > .5$. The second column contains comparable dates from the NBER. This reproduces Hamilton's (1989, table II) impressive result with a slightly different data set and shows that the ".5 rule" fairly accurately mimics the classical cycles of traditional cycle-dating methods. Note that the MS(4) model is even able to pick up the "double-dip" recession of 1980–1982.

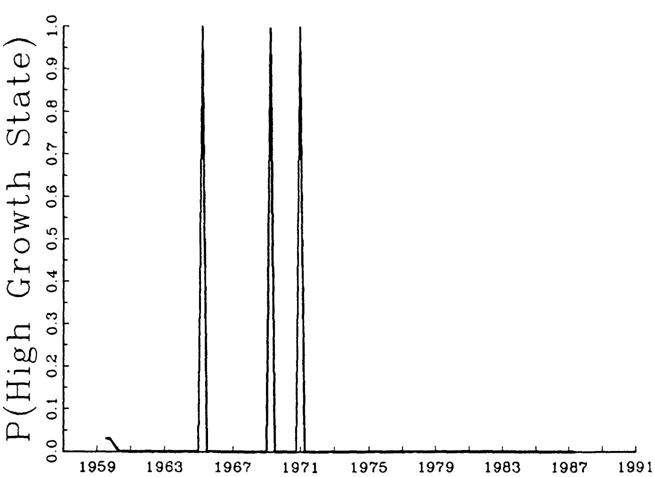

Figure 7. Time Path of Conditional Probabilities of the High-Growth State Calculated From the Estimates of Table 1 for Italy.

Mintz (1969) applied the NBER methodology to postwar Germany up to 1967. Classical business-cycle dates from her study are presented in Table 3 next to the MS(4) dates for Germany. The recession of 1961–1963 is matched perfectly by the Markov model using the .5 rule. But Mintz estimated the 1966–1967 recession to be three quarters longer than does the Markov model.

Table 3 also contains peak and trough dates from Phillips (1991) using the .5 rule. He specified a two-country Markov-switching model with an AR(1) component for each country. He pairs U.S. industrial production with Canada, the United Kingdom, and Germany. Compared to the MS(4) models here, his vector MS(1) models produce filtered inferences with both more and less extensive dynamics. The inferences have more dynamic information in that the transmission of business cycles from the United States to Germany, the United Kingdom, and Canada, and vice versa, are incorporated. The AR component has only one quarterly lag, however. Compared to the MS(4) model, the Phillips model generally produces more frequent and shorter lived cycles. Since the lead-lag relationships between countries are found to be weak, the difference in cycles is most likely due to the difference in AR lag length.

In summary, the MS(4) model matches the classical business-cycle dates of the NBER and Mintz (1969) quite closely using a .5 rule. At a minimum, the Markov-switching methodology appears to be a useful cross-check on traditional business-cycle dating methods. Since they are derived from very different premises, when

Table 3. Peak and Trough Dates for Seven Economies

	United States			United Kingdom		Germany			Japan	Canada		Switzerland	France
	MS(4)	NBER	Phillips	MS(4)	Phillips	MS(4)	Mintz	Phillips	MS(4)	MS(4)	Phillips	MS(4)	MS(4)
Peak				1957:4									
Trough				1958:2									
Peak	1960:1	1960:2											
Trough	1960:4	1961:1											
Peak						1961:1	1961:1						
Trough						1963:1	1963:1						
Peak								1963:2					
Trough								1963:3					
Peak								1964:2					
Trough								1964:3					
Peak						1966:2	1965:4	1966:3					
Trough						1967:1	1967:2	1967:1					
Peak													1967:4
Trough													1968:2
Peak	1969:3	1969:4											
Trough	1970:4	1970:4											
Peak					1971:4								
Trough					1972:1								
Peak	1973:4	1973:4	1973:4	1973:4	1973:3	1973:4		1974:1	1973:2				
Trough			1974:1		1974:1				1974:1				
Peak			1974:4		1974:3						1974:4	1974:2	1974:3
Trough	1975:1	1975:1	1975:1	1975:2	1975:2	1975:2		1975:2			1975:1	1976:1	1975:1
Peak								1977:1					
Trough								1977:2					
Peak	1980:1	1980:1	1980:3	1979:2	1979:4	1980:1		1980:1			1980:1		
Trough	1980:3	1980:3	1980:4	1981:2	1980:3			1980:4			1980:2		
Peak	1981:1	1981:3	1981:4		1982:3			1981:3		1981:2	1981:2		
Trough	1982:4	1982:4	1982:4		1982:4	1982:4		1983:2		1982:3	1982:4		
Peak					1985:2								
Trough					1988:3								

both approaches agree it is mutually reinforcing evidence. A conundrum arises, however, when the dates differ. It is not clear whether a disagreement should be taken as evidence against the NBER methods or the Markov model. Perhaps the two methods can be used in a complementary manner, each indicating directions for improvement in the other.

3. ARE BUSINESS CYCLES ASYMMETRIC?

Implicit in much of the research on business cycles going back to Keynes and before is the notion that they can be characterized as exhibiting sharp drops during contractions followed by gradual movements during expansions. A closely related idea is that contractions have shorter durations than expansions.

Conditional on being in the high-growth, or expansionary, state, the expected duration of an expansion is

$$\sum_{k=1}^{\infty} kp^{k-1}(1-p) \rightarrow (1-p)^{-1}. \tag{9}$$

Likewise, the expected duration of a recession is

$$\sum_{k=1}^{\infty} kq^{k-1}(1-q) \rightarrow (1-q)^{-1}. \tag{10}$$

The calculations are presented in Table 4. Country by country, the expected duration of expansions is uniformly greater than the expected duration of recessions across the seven economies.

An LR test formally comparing the transition prob-

Table 4. Tests of Asymmetry in Business Cycles

	U.S.	U.K.	Germany	Japan	Canada	Switzerland	France
				H_0: $p = q$			
LR test	4.75	4.76	1.72	4.09	5.72	5.21	6.98
p value	.029	.029	.19	.043	.017	.023	.0082
			Expected duration of expansions (quarters)				
$(1-p)^{-1}$	17.2	26.2	11.2	52.5	80.1	50.6	36.9
			Expected duration of recessions (quarters)				
$(1-q)^{-1}$	3.7	4.8	3.5	2.8	3.7	4.9	2.0

abilities is calculated for each of the seven economies using the estimated transition probabilities. LR statistics in Table 4 reject the null of $p = q$ for a test of size 5% for all countries except Germany.

4. SUMMARY AND CONCLUDING REMARKS

This article presents a thorough specification analysis of the Hamilton model of Markov switching for eight market economies. The analysis suggests only marginal improvements over a linear model. There is also evidence that Markov switching does not capture all nonlinearities for some economies. A further difficulty with the model is illustrated by the case of Italy, in which the likelihood identifies several outliers as a separate "state."

The specification analysis shows that the basic MS(4) model needs improvement if it is going to be worth the effort of departing from linear models. Hansen (1992) showed that a modification of the model to allow one of the AR parameters to shift between regimes in addition to the mean does fit U.S. GNP data better than either an MS(4) or an AR(4) model. In recent work, Hamilton and Susmel (1992) showed that extensions to multiple states can reduce the impact of outliers by, in effect, creating a separate outlier state.

The most innovative aspect of the Hamilton model is the ability to objectively date business cycles. Filtered and smoothed conditional probabilities indicate peak and trough dates for the United States that closely correlate with NBER dates. The cycle dates of Mintz (1969) for Germany are also closely reproduced by the MS(4) model. Using the filtered inferences to examine the international transmission of business cycles would be a useful extension of this article.

ACKNOWLEDGMENTS

I thank James Hamilton, Bruce Hansen, and Kerk Phillips for generously providing me with their data and computer code. I also thank the editor, an associate editor, and two anonymous referees for their very constructive suggestions.

[Received October 1991. Revised February 1993.]

REFERENCES

Beveridge, S., and Nelson, C. (1981), "A New Approach to Decompositions 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.
Brock, W., and Sayers, C. (1988), "Is the Business Cycle Characterized by Deterministic Chaos?" *Journal of Monetary Economics*, 22, 71–80.
Campbell, J. Y., and Mankiw, N. G. (1987), "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 102, 857–880.
Clark, P. K. (1987), "The Cyclical Component of U.S. Economic Activity," *Quarterly Journal of Economics*, 102, 797–814.
Diebold, F., and Rudebusch, G. (1990), "A Nonparametric Investigation of Duration Dependence in the American Business Cycle," *Journal of Political Economy*, 98, 596–616.
Engle, C., and Hamilton, J. (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review*, 80, 689–713.
Engle, R., Lilien, D., and Robins, R. (1987), "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 55, 391–407.
Hamilton, J. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.
——— (1991), "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions," *Journal of Business & Economic Statistics*, 9, 27–39.
——— (in press), "Estimation, Inference, and Forecasting of Time Series Subject to Changes in Regime," in *Handbook of Statistics–Volume 10*, eds. C. R. Rao and G. S. Maddala, New York: North-Holland.
Hamilton, J., and Susmel, R. (1992), "Autoregressive Conditional Heteroscedasticity and Changes in Regime," mimeo, University of California, San Diego.
Hansen, B. (1992), "The Likelihood Ratio Test Under Non-Standard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, S61–S82.
Harvey, A. C. (1985), "Trends and Cycles in Macroeconomic Time Series," *Journal of Business & Economic Statistics*, 3, 216–227.
Hinich, M., and Patterson, D. (1985), "Evidence of Nonlinearity in Daily Stock Returns," *Journal of Business & Economic Statistics*, 3, 69–77.
King, R. G., Plosser, C. I., Stock, J. H., and Watson, M. W. (1991), "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81, 819–840.
Kiviet, J. (1986), "On the Rigour of Some Misspecification Tests for Modelling Dynamic Relationships," *Review of Economic Studies*, 53, 241–61.
Koopmans, T. (1965), "Measurement Without Theory," in *Readings in Business Cycles*, eds R. Gordon and L. Klein, Homewood, IL: Richard Irwin, pp. 186–203.
Lam, P. (1990), "The Hamilton Model With a General Autoregressive Component," *Journal of Monetary Economics*, 26, 409–432.
Mintz, I. (1969), *Dating Postwar Business Cycles: Methods and Their Application to Western Germany, 1950–1967*, Chicago: National Bureau of Economic Research.
Mitchell, W., and Burns, A. (1938), *Statistical Indicators of Cyclical Revivals*, Chicago: National Bureau of Economic Research.
Neftci, S. (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, C., and Plosser, C. (1982), "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.
Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.
Phillips, K. (1991), "A Two-Country Model of Stochastic Output With Changes in Regime," *Journal of International Economics*, 31, 121–142.
Sichel, D. (1993), "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry*, 31, 224–236.
Stock, J. (1987), "Measuring Business Cycle Time," *Journal of Political Economy*, 95, 1240–1261.
Stock, J., and Watson, M. (1991), "A Probability Model of the Coincident Economic Indicators," in *Leading Indicators: New Approaches and Forecasting Records*, eds. K. Lahiri and G. Moore, New York: Cambridge Press, pp. 63–89.
Watson, M. W. (1986), "Univariate Detrending Methods With Stochastic Trends," *Journal of Monetary Economics*, 18, 49–76.