MEASURING BUSINESS CYCLES: A MODERN PERSPECTIVE

Francis X. Diebold and Glenn D. Rudebusch*

Abstract—In the first half of this century, special attention was given to two features of the business cycle: the comovement of many individual economic series and the different behavior of the economy during expansions and contractions. Recent theoretical and empirical research has revived interest in each attribute separately, and we survey this work. Notable empirical contributions are dynamic factor models that have a single common macroeconomic factor and nonlinear regime-switching models of a macroeconomic aggregate. We conduct an empirical synthesis that incorporates both of these features.

I. Introduction

It is desirable to know the facts before attempting to explain them; hence, the attractiveness of organizing business-cycle regularities within a model-free framework. During the first half of this century, much research was devoted to obtaining just such an empirical characterization of the business cycle. The most prominent example of this work was Burns and Mitchell (1946), whose summary empirical definition was:

Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle. (p. 3)

Burns and Mitchell’s definition of business cycles has two key features. The first is the comovement among individual economic variables. Indeed, the comovement among series, taking into account possible leads and lags in timing, was the centerpiece of Burns and Mitchell’s methodology. In their analysis, Burns and Mitchell considered the historical concordance of hundreds of series, including those measuring commodity output, income, prices, interest rates, banking transactions, and transportation services. They used the clusters of turning points in these individual series to determine the monthly dates of the turning points in the overall business cycle. Similarly, the early emphasis on the consistent pattern of comovement among various variables over the business cycle led directly to the creation of composite leading, coincident, and lagging indexes (e.g., Shishkin, 1961).

The second prominent element of Burns and Mitchell’s definition of business cycles is their division of business cycles into separate phases or regimes. Their analysis, as was typical at the time, treats expansions separately from contractions. For example, certain series are classified as leading or lagging indicators of the cycle, depending on the general state of business conditions.

Both of the features highlighted by Burns and Mitchell as key attributes of business cycles were less emphasized in postwar business-cycle models—particularly in empirical models where the focus was on the time-series properties of the cycle. Most subsequent econometric work on business cycles followed Tinbergen (1939) in using the linear difference equation as the instrument of analysis. This empirical work has generally focused on the time-series properties of just one or a few macroeconomic aggregates, ignoring the pervasive comovement stressed by Burns and Mitchell. Likewise, the linear structure imposed eliminated consideration of any nonlinearity of business cycles that would require separate analyses of expansions and contractions.

Recently, however, empirical research has revived consideration of each of the attributes highlighted by Burns and Mitchell. Notably, Stock and Watson (1989, 1991, 1993) have used a dynamic factor model to capture comovement by obtaining a single common factor from a set of many macroeconomic series, and Hamilton (1989) has estimated a nonlinear model for real GNP with discrete regime switching between periods of expansion and contraction.

This paper is part survey, part interpretation, and part new contribution. We describe the dynamic-factor and regime-switching models in some detail in sections II and III, and we sketch their links to recent developments in macroeconomics in section IV. The modern dynamic-factor and regime-switching literatures, however, have generally considered the comovement and regime-switching aspects of the business cycle in isolation of each other. We view that as unfortunate, as scholars of the cycle have simultaneously used both ideas for many decades. Thus, in section V, we attempt an empirical synthesis in a comprehensive framework that incorporates both factor structure and regime switching. We conclude in section VI.

II. Comovement: Factor Structure

In a famous essay, Lucas (1976) drew attention to a key business-cycle fact: outputs of broadly-defined sectors move together. Lucas’ view is part of a long tradition that
has stressed the coordination of activity among various economic actors and the resulting comovement in sectoral outputs.

Analysis of comovement in dynamic settings typically makes use of two nonparametric tools, the autocorrelation function and the spectral density function. In the time domain, one examines multivariate dynamics via the autocorrelation function, which gives the correlations of each variable with its own past and with the past of all other variables in the system. Such analyses are now done routinely, as in Backus and Kehoe (1992), who characterize the dynamics of output, consumption, investment, government purchases, net exports, money, and prices across ten countries and a hundred years.

Alternatively, one examines dynamics in the frequency domain via the spectral density function, the Fourier transform of the autocovariance function, which presents the same dynamic information but in a complementary fashion. The spectral density matrix decomposes variation and covariation among variables by frequency, permitting one to concentrate on the dynamics of interest (business-cycle dynamics, for example, correspond to periods of roughly 2–8 years). Transformations of both the real and imaginary parts of the spectral density matrix have immediate interpretation in business-cycle analysis; the coherence between any two economic time series effectively charts the strength of their correlation by frequency, while the phase charts lead/lag relationships by frequency. A good example of business-cycle analysis in the frequency domain is Sargent (1987), who examines the spectral density matrix of seven U.S. data series: real GNP, the unemployment rate, the interest rate, the change in real money stock, inflation, productivity, and real wages.2

Of course, one can analyze business-cycle data parametrically as well, by approximating the dynamic relationships with a particular statistical model. In this regard, the vector autoregression, introduced by Sims (1980), is ubiquitous. The moving-average representation (that is, the impulse-response function) of a vector autoregression of a set of macroeconomic variables provides a readily-interpretable characterization of dynamics, by charting the response of each variable to shocks to itself and the other variables.

Unfortunately, a vector-autoregressive study that attempts to capture the pervasive comovement among hundreds of series emphasized by Burns and Mitchell requires more degrees of freedom than are available in macroeconomic samples. Recent work provides crucial dimensionality reduction, however, because the dynamic comovements among large sets of macroeconomic variables are often well-described by a particular configuration of the vector autoregression associated with index structure, or factor structure.

Factor models have a long history of use in cross-sectional settings, and their generalization to dynamic environments is due to Sargent and Sims (1977), Geweke (1977) and Watson and Engle (1983). Important recent contributions include Stock and Watson (1989, 1991, 1993) and Quah and Sargent (1993), among others. The idea is simply that the comovement of contemporaneous economic variables may be due to the fact that they are driven in part by common shocks. In a one-factor model, for example, the behavior of the set of N variables is qualitatively similar to the behavior of just one variable, the common factor. This allows parsimonious modeling while nevertheless maintaining fidelity to the notion of pervasive macroeconomic comovement.3

Let us focus on the dynamic factor model of Stock and Watson (1991), which was developed as a modern statistical framework for computing a composite index of coincident indicators. In their one-factor model, movements in the N macroeconomic variables of interest, $x_t$, are determined by changes in the one-dimensional unobserved common factor, $f_t$, and by the N-dimensional idiosyncratic component, $u_t$:

$$x_t = \beta + \lambda f_t + u_t$$

$$N \times 1 \quad N \times 1 \quad N \times 1 \quad 1 \times 1 \quad N \times 1$$

$$D(L) u_t = \epsilon_t$$

$$N \times N \quad N \times 1 \quad N \times 1$$

$$\phi(L) (f_t - \delta) = \eta_t$$

$$1 \times 1 \quad 1 \times 1 \quad 1 \times 1$$

(1993), among others. The idea is simply that the comovement of contemporaneous economic variables may be due to the fact that they are driven in part by common shocks. In a one-factor model, for example, the behavior of the set of N variables is qualitatively similar to the behavior of just one variable, the common factor. This allows parsimonious modeling while nevertheless maintaining fidelity to the notion of pervasive macroeconomic comovement.3

Let us focus on the dynamic factor model of Stock and Watson (1991), which was developed as a modern statistical framework for computing a composite index of coincident indicators. In their one-factor model, movements in the N macroeconomic variables of interest, $x_t$, are determined by changes in the one-dimensional unobserved common factor, $f_t$, and by the N-dimensional idiosyncratic component, $u_t$:

$$x_t = \beta + \lambda f_t + u_t$$

$$N \times 1 \quad N \times 1 \quad N \times 1 \quad 1 \times 1 \quad N \times 1$$

$$D(L) u_t = \epsilon_t$$

$$N \times N \quad N \times 1 \quad N \times 1$$

$$\phi(L) (f_t - \delta) = \eta_t$$

$$1 \times 1 \quad 1 \times 1 \quad 1 \times 1$$

All idiosyncratic stochastic dynamics are driven by $\epsilon_t$ while all common stochastic dynamics, which are embodied in the common factor, are driven by $\eta_t$. Identification may be achieved in many ways. Stock and Watson, for example, impose (1) orthogonality at all leads and lags of $\{u_{1t}, \ldots, u_{Nt}, f_t\}$ (which is achieved by making $D(L)$ diagonal and $\{f_{1t}, \ldots, f_{Nt}\}$ orthogonal at all leads and lags), and (2) var($\eta_t$) = 1.

III. Nonlinearity: Regime Switching

Underlying much of the traditional business-cycle literature is the notion that a good statistical characterization of business-cycle dynamics may require some notion of regime switching between “good” and “bad” states.4 Models incorporating regime switching have a long tradition in dynamic econometrics.5 One recent time-series model

---

2 In the frequency domain, Sargent (1987, p. 282) offers the following update of Burns and Mitchell’s definition: “ . . . the business cycle is the phenomenon of a number of important economic aggregates (such as GNP, unemployment, and layoffs) being characterized by high pairwise coherences at the low business cycle frequencies.”

3 It is interesting to note that parallel structures may exist in many financial markets, which makes sense to the extent that asset prices accurately reflect fundamentals, which themselves have factor structure. See Singleton (1980), Bollerslev, Engle and Woolridge (1988), and Diebold and Nerlove (1989), among others, for examples of factor structure in both the conditional means and conditional variances of various asset returns.

4 Again, parallel structures may exist in financial markets. Regime switching has been found, for example, in the conditional mean dynamics of interest rates (Hamilton, 1988; Cecchetti, Lam and Mark, 1990) and exchange rates (Engel and Hamilton, 1990), and in the conditional variance dynamics of stock returns (Hamilton and Susmel, 1994).

5 Key early contributions include the early work of Quandt (1958) and Goldfeld and Quandt (1973).
squarely in line with the regime-switching tradition is the “threshold” model (e.g., Tong, 1983; Potter, 1995). In a threshold model, the regime switches according to the observable past history of the system.

Although threshold models are of interest, models with latent states as opposed to observed states may be more appropriate for business-cycle modeling. Hamilton (1989, 1990, 1994) develops such models. In Hamilton’s regime-switching setup, time-series dynamics are governed by a finite-dimensional parameter vector that switches (potentially each period) depending upon which of two unobservable states is realized, with state transitions governed by a first-order Markov process.

To make matters concrete, let’s take a simple example. Let \( \{s_t\}_{t=1}^T \) be the (latent) sample path of a two-state first-order Markov process, taking values 0 or 1, with transition probability matrix given by

\[
M = \begin{bmatrix}
P_{00} & 1 - P_{00} \\
1 - P_{11} & P_{11}
\end{bmatrix}.
\]

Let \( \{y_t\}_{t=1}^T \) be the sample path of an observed time series that depends on \( \{s_t\}_{t=1}^T \) such that the density of \( y_t \) conditional upon \( s_t \) is

\[
f(y_t|s_t; \theta) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{(y_t - \mu_{s_t})^2}{2\sigma^2} \right).
\]

Thus, \( y_t \) is Gaussian white noise with a potentially switching mean. The two means around which \( y_t \) moves are of particular interest, and for example, correspond to episodes of differing growth rates ("expansions" and "contractions").

The central idea of regime switching is simply that expansions and contractions may be usefully treated as different probabilistic objects. This idea has been an essential part of the Burns-Mitchell-NBER tradition of business-cycle analysis and is also manifest in the great interest in the popular press, for example, in identifying and predicting turning points in economic activity. Yet it is only within a regime-switching framework that the concept of a turning point has intrinsic meaning. Recent contributions that have emphasized the use of probabilistic models in the construction and evaluation of turning-point forecasts and chronologies include Neftci (1984) and Diebold and Rudebusch (1989).

Various seemingly-disparate contributions may be readily interpreted within the context of the basic switching model. One example is Neftci’s (1984) well-known analysis of business-cycle asymmetry, which amounts to asking whether the transition probability matrix is symmetric. Another example is Potter’s (1995) and Sichel’s (1994) evidence for the existence of a “recovery” regime of very fast growth at the beginning of expansions, which corresponds to a “third state” in business-cycle dynamics.

Yet another class of examples concerns recent analyses of business-cycle duration dependence, which amount to asking whether the transition probabilities vary with length-to-date of the current regime. Diebold and Rudebusch (1990), Diebold, Rudebusch, and Sichel (1993), and Filardo (1994) have found positive duration dependence in postwar U.S. contractions; that is, the longer a contraction persists, the more likely it is to end soon. Similar results have been obtained by Durland and McCurdy (1994) using the technology of semi-Markov processes. Other forms of time-variation in business-cycle transition probabilities may be important as well. Ghysels (1993, 1994), in particular, argues that business-cycle transition probabilities vary seasonally and provides formal methods for analyzing such variation.

The business-cycle duration dependence literature highlights the fact that economic considerations may suggest the potential desirability of allowing the transition probabilities to vary through time. The duration dependence literature focuses on trend and seasonal variation in transition probabilities, but in certain contexts it may be desirable to allow for more general time-variation, as in Diebold, Lee, and Weinbach (1994) and Filardo (1994), who let the transition probabilities evolve as logistic functions of exogenous variables, \( z_t \):

\[
M_t = \begin{bmatrix}
\frac{\exp(z_t' \beta_0)}{1 + \exp(z_t' \beta_0)} & 1 - \frac{\exp(z_t' \beta_0)}{1 + \exp(z_t' \beta_0)} \\
\frac{\exp(z_t' \beta_1)}{1 + \exp(z_t' \beta_1)} & \frac{\exp(z_t' \beta_1)}{1 + \exp(z_t' \beta_1)}
\end{bmatrix}.
\]

IV. Factor Structure and Regime Switching: Links to Macroeconomic Theory

In this section, as further motivation, we describe some of the links between macroeconomic theory and factor structure and regime switching. We use convex equilibrium business-cycle models to motivate the appearance of factor structure and non-convex models with multiple equilibria to motivate regime switching; however, we hasten to add that these pairings are by no means exclusive. Moreover, of course, our ultimate interest lies in models that simultaneously display factor structure and regime-switching behavior, which as the following discussion suggests, might occur in a variety of ways.

A. Macroeconomic Theory and Factor Structure

The econometric tradition of comovement through factor structure is consistent with a variety of modern dynamic
macroeconomic models. Here we highlight just one—a linear-quadratic equilibrium model—in order to motivate the appearance of factor structure. We follow the basic setup of Hansen and Sargent (1993), which although arguably rigid in some respects, has two very convenient properties. First, the discounted dynamic programming problem associated with the model may be solved easily and exactly. Second, the equilibria of such models are precisely linear (that is, precisely a vector autoregression), thereby bringing theory into close contact with econometrics.

Preferences are quadratic and are defined over consumption of services, \( s_t \), and work effort, \( l_t \), with preference shocks, \( b_t \), determining a stochastic bliss point. There are four linear constraints on the utility maximization. The first represents the linear technology: a weighted average of the output of consumption goods, \( c_t \), intermediate goods, \( g_t \), and investment goods, \( i_t \), equals a linear combination of lagged capital stock, \( k_{t-1} \), and work effort, plus the technology shock, \( d_t \). The second is the law of motion for the capital stock: Capital accumulates through additional net investment. The third is the law of motion for “household capital,” \( h_t \), which is driven by consumption expenditures. The last specifies that current consumption services depend on both lagged household capital and current consumption.

Formally, the planning problem associated with this model is

\[
\max \frac{1}{2} E \sum_{t=0}^{\infty} \beta^t \left[ (s_t - b_t) \cdot (s_t - b_t) + I_t^2 \right]
\]

subject to the four constraints

\[
\begin{align*}
\alpha_1 c_t + \alpha_2 g_t + \alpha_3 i_t &= \alpha_4 k_{t-1} + \alpha_5 l_t + d_t \\
k_t &= \beta_1 k_{t-1} + \beta_2 l_t \\
h_t &= \gamma_1 h_{t-1} + \gamma_2 c_t \\
s_t &= \delta_1 h_{t-1} + \delta_2 c_t.
\end{align*}
\]

The exogenous uncertainty \( (e_{t+1}) \) in the model evolves according to

\[
e_{t+1} = \rho e_t + w_{t+1}
\]

where \( w_{t+1} \) is zero-mean white noise. The preference and technology shocks \( (b_t \) and \( d_t) \) are linear transformations of the \( e_t \),

\[
b_t = U_b e_t \\
d_t = U_d e_t.
\]

Importantly for our purposes, note that this framework can potentially describe the determination of a large set of series. All variables (except \( l_t \)) can be considered as vectors of different goods or services with the parameters interpreted as conformable matrices.

The equilibrium of this economy is a linear stochastic process and can be presented by a vector autoregression constrained by cross-equation restrictions, with state-space form

\[
\begin{align*}
\alpha_{t+1} &= A \alpha_t + C w_{t+1} \\
\alpha_t &= G \alpha_{t} + v_t,
\end{align*}
\]

where the state vector \( \alpha_t \) contains \( h_t, k_t, \) and \( e_t \). \( \alpha_t \) contains any variable that can be expressed as a linear function of the state variables. Note that this vector autoregression will be singular so long as the number of shocks is less than the number of variables in the system. Fewer shocks than observables is the rule in economic models. The standard setups have just a few preference and technology shocks driving a comparatively large number of decision variables, thereby building in singularity. In fact, in the leading case of a single technology shock and no preference shocks, one shock is responsible for all variation in the choice variables, resulting in an equilibrium that maps into a special (singular) case of the one-factor model discussed earlier. In that special case, there are no idiosyncratic shocks (or equivalently, they have zero variance).

To reconcile the singular equilibrium from the model economy with the clearly non-singular nature of the data, measurement error is often introduced. The state-space representation becomes

\[
\begin{align*}
\alpha_{t+1} &= A \alpha_t + C w_{t+1} \\
\alpha_t &= G \alpha_{t} + v_t,
\end{align*}
\]

where \( v_t \) is a martingale-difference sequence. In single-shock linear-quadratic models with measurement error, the equilibria are precisely of the single-factor form, with non-degenerate idiosyncratic effects.

Feeling constrained by linear technology and quadratic preferences, many authors have recently focused on models that are not linear-quadratic. The formulation is basically the same as in the linear-quadratic case, but the mechanics are more complicated. The discounted dynamic programming problem associated with the recursive competitive equilibrium can only be solved approximately; however, the decision rules are nevertheless well-approximated linearly near the steady state. Under regularity conditions, the equilibrium is a Markov process in the state variables, and if that Markov process converges to an invariant distribution, then a vector-autoregressive representation exists. Again, the vector autoregression is only an approximation to the generally nonlinear decision rules, and its computation can be tedious. However, the availability of a factor structure for modelling this approximation remains.

---

8 Consumption appears in both of the last two equations in order to capture both its durable and nondurable aspects.

9 See Sargent (1989) and Hansen and Sargent (1993), among others.

B. Macroeconomic Theory and Regime Switching

Regime-switching behavior is also consistent with a variety of macroeconomic models. Here we focus on models with coordination failures, which produce multiple equilibria. In what follows, we shall provide a brief overview of this theoretical literature and its relation to the regime-switching model.

Much has been made of the role of spillovers and strategic complementarities in macroeconomics (Cooper and John, 1988). “Spillover” simply refers to a situation in which others’ strategies affect one’s own payoff. “Strategic complementarity” refers to a situation in which others’ strategies affect one’s own optimal strategy. Spillovers and strategic complementarities can produce a variety of macroeconomic models. Here we focus on models of optimal search (e.g., Diamond, 1982), where thick-market externalities ensure that the likelihood of successful search depends on the intensity of search undertaken by others, which in turn affects one’s own optimal search intensity. In short, search is more desirable when other agents are also searching, because it is likely to be more productive.

Spillovers and strategic complementarities may have important macroeconomic effects. For example, the appearance of aggregate increasing returns to scale (e.g., Hall, 1991) may simply be an artifact of the positive externalities associated with high output levels in the presence of spillovers and strategic complementarities rather than true increasing returns in firms’ technologies. Indeed, Caballero and Lyons (1992) find little evidence of increasing returns at the individual level, yet substantial evidence at the aggregate level, suggesting the importance of spillovers and strategic complementarities.

Spillovers and strategic complementarities can produce multiple equilibria, the dynamics of which may be well-approximated by statistical models involving regime switching. In fact, Cooper and John (1988) stress the existence of multiple equilibria, with no coordination mechanism, as a common theme in a variety of seemingly-unrelated models displaying spillovers and strategic complementarities. Moreover, the equilibria are frequently pareto-rankable. Situations arise, for example, in which an economy is in a low-output equilibrium such that all agents would be better off at higher output levels, but there is no coordination device to facilitate the change.

Recent work has provided some mechanisms for endogenizing switches between equilibria. One approach involves variations on Keynesian “animal spirits,” or self-fulfilling waves of optimism and pessimism, as formalized by Azariadis (1981) and Cass and Shell (1983). Notably, Diamond and Fudenberg (1989) demonstrate in a search framework the existence of rational-expectations sunspot equilibria in which agents’ beliefs about cycles are self-fulfilling. Howit and McAfee (1992) obtain results even more in line with our thesis in a model in which waves of optimism and pessimism evolve according to a Markov process. The statistical properties of equilibria from their model are well-characterized by a Markov regime-switching process.

Finally, Cooper (1994) proposes a history-dependent selection criterion in an economy with multiple Nash equilibria corresponding to different levels of productivity. The Cooper criterion reflects the idea that history may create a focal point: a person’s recent experience is likely to influence her expectations of others’ future strategic behavior, resulting in a slow evolution of conjectures about others’ actions. Cooper’s analysis highlights the importance of learning to respond optimally to the strategic actions of others. The Cooper criterion leads to persistence in the equilibrium selected, with switching occurring as a consequence of large shocks, phenomena which again may be well-characterized by statistical models involving regime switching.

Other history-dependent theoretical models have been proposed by Startz (1994) and Acemoglu and Scott (1993). These include the same “learning-by-doing” dynamic externality that drives the “new growth theory” models. Again, shocks cause endogenous switching between high-growth and low-growth states.

V. Synthesis: Regime Switching in a Dynamic Factor Model

We have argued that both comovement through factor structure and nonlinearity through regime switching are important elements to be considered in an analysis of business cycles. It is unfortunate, therefore, that the two have recently been considered largely in isolation from each other. In what follows, we sketch a framework for the analysis of business-cycle data that incorporates both factor structure and regime switching in a natural way. This framework, although not formally used before, may be a good approximation to the one implicitly adopted by many scholars of the cycle.

A. A Prototypical Model

Consider a dynamic factor model in which the factor switches regimes. First consider a switching model for the factor $f_t$; we work with a slightly richer model than before, allowing for $p$th-order autoregressive dynamics. Again let $\{s_t\}_{t=1}^T$ be the sample path of a latent Markov process, taking on values 0 and 1, let $\{f_{t,j}\}_{j=1}^N$ be the sample path of the factor (which depends on $\{s_t\}_{t=1}^T$), and collect the relevant history of the factor and state in the vector $h_t = (s_{t-1}, \ldots, s_{t-p}, f_{t-1}, \ldots, f_{t-p})$. The probabilistic dependence of $f_t$ on $h_t$ is summarized by the conditional density,
The latent factor, then, follows a \( p \)-th order Gaussian autoregression with potentially changing mean. The two means around which the factor moves are of particular interest; call them \( \mu_0 \) (slow growth) and \( \mu_1 \) (fast growth).

We then assemble the rest of the model around the regime-switching factor. We write

\[
\Delta x_t = \beta + \lambda f_t + u_t
\]

\[
N \times 1 \quad N \times 1 \quad 1 \times 1 \quad N \times 1
\]

\[
D(L) \quad u_t = \varepsilon_t
\]

\[
N \times N \quad N \times 1 \quad N \times 1
\]

as earlier. In Frischian terms, the model as written has a regime-switching "impulse" with a stable "propagation mechanism." Many variations on the theme of this basic setup are of course possible.

### B. A Look at the Data

Let us first describe the data. We examine quarterly economic indicators, 1952.1–1993.1, as described in detail in table 1. The data include three composite indexes of coincident indicators, corresponding to three alternative methodologies: Commerce Department, modified Commerce Department, and Stock-Watson. The component indicators underlying the Commerce Department and modified Commerce Department indexes are identical (personal income less transfer payments, index of industrial production, manufacturing and trade sales, and employees on non-agricultural payrolls); only their processing differs slightly (see Green and Beckman, 1992). The Stock-Watson index introduces a change in the list of underlying indicators (employees on nonagricultural payrolls is replaced by hours of employees on nonagricultural payrolls) and processes the underlying component indicators differently than either the Commerce Department or modified Commerce Department indexes. We obtained qualitatively similar results from all of the indexes; thus, we shall focus here on the Commerce Department’s modified Composite Coincident Index. Henceforth, we shall refer to it simply as “the Composite Coincident Index.”

We graph the log of the Composite Coincident Index in figure 1. It tracks the business cycle well, with obvious and pronounced drops corresponding to the NBER-designated recessions of 1958, 1960, 1970, 1974, 1980, 1982 and 1990. We similarly graph the logs of the four components of the Composite Coincident Index in figure 2.\(^{14}\) Their behavior closely follows that of the Composite Coincident Index; in particular, there seems to be commonality among switch times.

We shall not provide maximum-likelihood estimates (or any other estimates) of a fully-specified dynamic-factor model with regime-switching factor. To do so would be premature at this point. Instead, we shall sift the data in two simple exercises to provide suggestive evidence as to whether the data accord with our basic thesis.

First, we work directly with the Composite Coincident Index, which is essentially an estimate of the common factor underlying aggregate economic activity.\(^ {15}\) We ask whether its dynamics are well-approximated by a Markov-

---

\(^{14}\) Each of the four component indicators is graphed on a different scale to enable their presentation in one graph. For this reason, no scale appears on the vertical axis of the graph.

\(^{15}\) Stock and Watson motivate and derive their index in precisely this way. The Commerce indexes are attempts at the same methodology, albeit less formally.
switching model. We fit a switching model to one hundred times the change in the natural logarithm of the Composite Coincident Index, with one autoregressive lag and a potentially switching mean.

The results appear in the second column of table 2. Several points are worth mentioning. First, the state-0 mean is significantly negative, and the state-1 mean is significantly positive, and the magnitudes of the estimates are in reasonable accord with our priors. Second, the within-state dynamics display substantial persistence. Third, the estimates of $P_{00}$ and $P_{11}$ accord with the well-known fact that expansion durations are longer than contraction durations on average. Fourth, the smoothed (that is, conditional upon all observations in the sample) probabilities that the Composite Coincident Index was in state 0 (graphed in figure 3) are in striking accord with the professional consensus as to the history of U.S. business cycles.\footnote{17}{They follow the NBER chronology closely, for example.}

In our second exercise, we fit switching models to the individual indicators underlying the Composite Coincident Index and examine the switch times for commonality. In a similar fashion to our analysis of the Composite Coincident Index, we fit models to one hundred times the change in the natural logarithm of each of the underlying coincident indicators, with one autoregressive lag and potentially switching means.

The results appear in columns three through six of table 2.\footnote{18}{Again, we use the startup values shown in the first column of table 2.}

The component-by-component results are qualitatively similar to those for the Composite Coincident Index, as would be expected in the presence of a regime-switching common factor. Further evidence in support of factor structure emerges in figure 4, in which we graph the time series of smoothed state-0 probabilities for each of the four component coincident indicators. There is commonality in

\begin{table}[h]
\centering
\caption{Estimated AR(1) Markov-Switching Models}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{START} & \textbf{CCIM} & \textbf{PITP} & \textbf{ENAP} & \textbf{IP} & \textbf{MTS} \\
\hline
$\mu_0$ & -0.50 & -0.91 & -0.75 & -0.54 & -4.12 & -2.26 \\
& (0.17) & (0.45) & (0.13) & (0.70) & (0.96) \\
$\mu_1$ & 0.50 & 0.97 & 0.88 & 0.61 & 1.16 & 1.01 \\
& (0.11) & (0.15) & (0.09) & (0.29) & (0.27) \\
$\phi_1$ & 0.40 & 0.66 & 0.35 & 0.97 & 0.52 & 0.38 \\
& (0.10) & (0.10) & (0.08) & (0.09) & (0.11) \\
$\sigma^2$ & 0.80 & 0.31 & 0.48 & 0.10 & 2.04 & 2.13 \\
& (0.04) & (0.08) & (0.01) & (0.24) & (0.38) \\
$P_{00}$ & 0.75 & 0.70 & 0.68 & 0.63 & 0.57 & 0.45 \\
& (0.10) & (0.16) & (0.11) & (0.15) & (0.28) \\
$P_{11}$ & 0.90 & 0.92 & 0.96 & 0.95 & 0.96 & 0.95 \\
& (0.02) & (0.03) & (0.02) & (0.02) & (0.04) \\
\hline
\end{tabular}
\end{table}

Notes: The column labeled "START" contains the startup values used for iteration. The other column labels denote the variable (defined in table 1) to which the Markov-switching model is fitted. Asymptotic standard errors appear in parentheses. The sample period is 1952.1-1993.1.
switch times, which is indicative of factor structure. Note, however, that the ability of the individual component indicators to track the business cycle (as captured in the smoothed state-0 probabilities for each of the component indicators) is inferior to that of the Composite Coincident Index. This is consistent with the switching-factor argument. Individual series are swamped by measurement error and hence provide only very noisy information on the state of the business cycle, but moving to a multivariate framework enables more precise tracking of the cycle.

C. Assessing Statistical Significance

Thus far, our empirical work has proceeded under the assumption of regime switching. It is also of obvious interest to test for regime switching—that is, to test the null hypothesis of one state against the alternative of two. The vast majority of the dozens of papers fitting Markov switching models make no attempt to test that key hypothesis. This is because the econometrics are nonstandard. Boldin (1990), Hansen (1992, 1996a, 1996b) and Garcia (1992) point out that the transition probabilities that govern the Markov switching are not identified under the one-state null, and moreover, that the score with respect to the mean parameter of interest is identically zero if the probability of staying in state 1 is either 0 or 1. In either event, the information matrix is singular.

Hansen proposes a bounds test that is valid in spite of these difficulties, but its computational difficulty has limited its applicability. A closely related approach, suggested by Garcia, is operational, however. The key is to treat the transition probabilities as nuisance parameters (ruling out from the start the problematic boundary values 0 and 1) and to exploit another of Hansen’s (1992) results, namely that the likelihood ratio test statistic for the null hypothesis of one state is the supremum over all admissible values of the nuisance parameters (the transition probabilities).

Let \( \theta = (\mu_0, \mu_1, \phi_0, \cdots, \phi_0, \sigma^2)^T \) be the set of all model parameters other than \( \rho_{00} \) and \( \rho_{11} \). We write the log likelihood as a function of three parameters (one vector parameter, and two scalar parameters),

\[
\ln L(\theta, \rho_{00}, \rho_{11}) = \ln P(y_1, \ldots, y_T; \theta, \rho_{00}, \rho_{11}).
\]

If we let a “hat” denote a maximum-likelihood estimator, then the maximized value of the log likelihood is \( \ln L(\hat{\theta}, \hat{\rho}_{00}, \hat{\rho}_{11}) \). Now consider maximizing the likelihood under the constraint (corresponding to the null hypothesis of one state) that \( \mu_0 = \mu_1 \). In that case, \( \rho_{00} \) and \( \rho_{11} \) are unidentified, so the maximized value of the log likelihood function is the same for any values of \( \rho_{00} \) and \( \rho_{11} \). Therefore, we simply write \( \ln L(\theta^*) \), where \( \theta^* \) is the constrained maximum-likelihood estimator of \( \theta \). Assembling all of this, we write the likelihood-ratio statistic for the null hypothesis of one state as

\[
LR = 2 \left[ \ln L(\hat{\theta}, \hat{\rho}_{00}, \hat{\rho}_{11}) - \ln L(\theta^*) \right].
\]

Now consider a different constrained likelihood maximization problem, in which we maximize the likelihood for arbitrary fixed values of \( \rho_{00} \) and \( \rho_{11} \). We denote the constrained maximum-likelihood estimator by \( \hat{\theta}(\rho_{00}, \rho_{11}) \); the resulting maximized value of the constrained log likelihood is \( \ln L(\hat{\theta}(\rho_{00}, \rho_{11}), \rho_{00}, \rho_{11}) \). Now form the likelihood-ratio statistic that compares the restrictions associated with \( \hat{\theta}(\rho_{00}, \rho_{11}) \) to those associated with \( \theta^* \), namely

\[
LR(\rho_{00}, \rho_{11}) = 2 \left[ \ln L(\hat{\theta}(\rho_{00}, \rho_{11}), \rho_{00}, \rho_{11}) - \ln L(\theta^*) \right].
\]

Garcia (1992), building on Hansen (1996a), establishes that

\[
LR = \sup_{\rho_{00}, \rho_{11}} LR(\rho_{00}, \rho_{11}),
\]

where \( \rho_{00} \) and \( \rho_{11} \) are restricted to the interior of the unit interval. This makes clear the intimate connection of this testing problem to Andrews’ (1993) test of structural change with breakpoint identified from the data, and not surprisingly, the limiting distribution of \( LR \) is of precisely the same form.20

Table 3 reports \( LR \) statistics calculated for the Composite index as well as its components. For the AR(1) case, which is the one relevant to the estimation results presented earlier,

---

19 See Hamilton’s (1994) survey, and the many papers cited there.

20 The results of Giné and Zinn (1990) and Stinchcombe and White (1993), used by Diebold and Chen (1996) to argue the validity of the bootstrap in Andrews’ (1993) case, are relevant here as well.
the asymptotic distribution of $LR$ has been characterized and tabulated by Garcia (1992) and shown to be accurate in samples of our size.\textsuperscript{21} Using the Garcia critical values, it is clear that the null hypothesis of no switching is overwhelmingly rejected for the CCI and each of its components.

A battery of diagnostic tests revealed that the AR(1) specification is typically quite good, although for some series (particularly manufacturing and trade sales) there is evidence that inclusion of a few more lags may improve the approximation. This raises the question of whether the AR(1) model is inducing serial correlation in the error, which is spuriously being picked up by the regime-switching dynamics. Thus, as a robustness check, we also present LR statistics for higher orders of autoregressive approximation in table 3. As in the AR(1) case, the asymptotic null distribution of LR depends on nuisance parameters (the autoregressive coefficients), but also as before, the dependence appears to be minor. Garcia, for example, calculates the 1% critical value for a particular AR(4) to be 11.60, which is little different than the AR(1) critical values. Each of our test statistics in table 3 is so much larger than the range of available critical values that even though they may not be strictly applicable, a strong rejection of the null hypothesis of one state appears unavoidable for the Composite Coincident Index as well as for all of its components.

### VI. Concluding Remarks and Directions for Future Research

We have argued that a model with factor structure and regime switching is a useful modern distillation of a long tradition in the analysis of business-cycle data. We proposed one stylized version of such a model, and we suggested its compatibility with macroeconomic data and macroeconomic theory.

Let us summarize our stance on the importance of the two attributes of the business cycle on which we have focused. It appears to us that comovement among business-cycle indicators is undeniable. This comovement could perhaps be captured by a VAR representation, if very long time series were available. The factor structure that we have advocated goes further, in that it implies restrictions on the VAR representation, restrictions that could be at odds with the data. Although more research is needed on that issue, the factor model is nothing more than a simple and parsimonious way of empirically implementing the common idea of fewer sources of uncertainty than variables.

The alleged nonlinearity of the business cycle is open to more dispute. The linear model has two key virtues: (1) it works very well much of the time, in economics as in all the sciences, in spite of the fact that there is no compelling a priori reason why it should, and (2) there is only one linear model, in contrast to the many varieties of nonlinearity. Why worry, then, about nonlinearity in general, and regime switching in particular?

First, a long tradition in macroeconomics, culminating with the earlier-discussed theories of strategic complementarities and spillovers in imperfectly competitive environments, thick-market externalities in search, self-fulfilling prophesies, and so on, makes predictions that seem to accord with the regime-switching idea.

Second, regime-switching models seem to provide a good fit to aggregate output data. Our rejections of the no-switching null hypothesis, in particular, appear very strong.

Third, the cost of ignoring regime switching, if in fact it occurs, may be large. Business people, for example, want to have the best assessments of current and likely future economic activity, and they are particularly concerned with turning points. Even tiny forecast improvements that may arise from recognizing regime switching may lead to large differences in profits. Similarly, for policy makers, if regime switching corresponds to movements between Pareto-rankable equilibria, there are important policy implications.\textsuperscript{22}

Fourth, macroeconomists, more generally, are interested in a host of issues impinged upon by the existence or nonexistence of regime switching. Optimal decision rules for consumption and investment (including inventory investment), for example, may switch with regime, as may agents' ability to borrow.

There are many directions for future research. The obvious extension is computation of full system estimates for the full dynamic-factor/Markov-switching model, which is straightforward conceptually but has been computationally infeasible thus far. Two avenues appear promising. One approach employs a multimove Gibbs sampler, in conjunction with a partially non-Gaussian state-space representation and a simulated EM algorithm, as developed recently by Shephard (1994) and de Jong and Shephard (1995). A similar approach from a Bayesian perspective is proposed in Kim (1994b).

A second approach involves using Kim's (1994a) filtering

---

\textsuperscript{21} The null distribution depends, even asymptotically, on the (unknown) true value of the autoregressive parameter. Fortunately, however, the dependence is slight; for example, Garcia's 1% critical values only vary from 11.54 to 11.95 over an autoregressive parameter range of -0.5 to 0.8.

\textsuperscript{22} Moreover, countercyclical policy may itself introduce nonlinearities if it is applied only in extreme situations. See Zarnowitz and Moore (1982) and Becketti and Haltiwanger (1987).
algorithm for a general class of models in state-space form, of which ours is a special case. The Kim algorithm maximizes an approximation to the likelihood rather than the exact likelihood, but the algorithm is fast and the approximation appears accurate. Presently, Chauvet (1995) is using the Kim algorithm to estimate the model and extract estimates of the factor (the "coincident index"), using both quarterly and monthly data and a variety of detrending procedures.

REFERENCES


Startz, Richard, "Growth States and Sectoral Shocks," manuscript, Department of Economics, University of Washington (1994).


