THE ASYMMETRIC BUSINESS CYCLE

James Morley and Jeremy Piger*

Abstract—The business cycle is a fundamental yet elusive concept in macroeconomics. In this paper, we consider the problem of measuring the business cycle. First, we argue for the output-gap view that the business cycle corresponds to transitory deviations in economic activity away from a permanent, or trend, level. Then we investigate the extent to which a general model-based approach to estimating trend and cycle for the U.S. economy leads to measures of the business cycle that reflect models versus the data. We find empirical support for a nonlinear time series model that produces a business cycle measure with an asymmetric shape across NBER expansion and recession phases. Specifically, this business cycle measure suggests that recessions are periods of relatively large and negative transitory fluctuations in output. However, several close competitors to the nonlinear model produce business cycle measures of widely differing shapes and magnitudes. Given this model-based uncertainty, we construct a model-averaged measure of the business cycle. This measure also displays an asymmetric shape and is closely related to other measures of economic slack such as the unemployment rate and capacity utilization.

I. Introduction

THE business cycle is a broad term that connotes the inherent fluctuations in economic activity. Research on the measurement of business cycles has a long tradition in macroeconomics, with an early example provided by Wesley Mitchell (1927), founder of the National Bureau of Economic Research (NBER). An integral part of business cycle measurement is its definition. Mitchell and the NBER defined the business cycle in terms of the alternation between periods of expansion and recession in the level of economic activity (which can be denoted the alternating-phases definition). One popular alternative definition is that the business cycle represents transitory fluctuations in economic activity away from a permanent, or "trend," level (which can be denoted the output-gap definition). This definition is associated with work on the U.S. business cycle by Beveridge and Nelson (1981), who propose an approach to measuring the business cycle based on long-horizon forecasts.

In this paper, we revisit the problem of measuring the business cycle. We begin by arguing for the output-gap notion of the business cycle as transitory fluctuations in economic activity. We then discuss how to conduct trend and cycle decomposition based on long-horizon forecasts for linear and nonlinear time series models, including how to implement an approach developed in Morley and Piger (2008) for empirically relevant regime-switching processes.

When we apply model-based trend and cycle decomposition to U.S. real GDP, we find that the estimated business cycle is highly dependent on model specification, with the key distinction being between linear models that imply symmetric fluctuations around trend and nonlinear regime-switching models that imply asymmetric deviations away from trend. In order to discriminate among the different measures of the business cycle, we use information criteria to evaluate the models and, in certain key cases, carry out formal hypothesis tests. The empirical results support a particular class of nonlinear regime-switching models and an asymmetric business cycle. However, the results also reveal several close competitors to the preferred nonlinear model that produce business cycle measures of widely differing shapes and magnitudes. This implies significant model-based uncertainty regarding the appropriate business cycle measure.

Given this uncertainty, we proceed to construct a model-averaged measure of the business cycle. In doing so, we are motivated by the principle of forecast combination, which is the idea that a combined forecast can be superior to all of the individual forecasts that go into its construction. For the weights used in combining model-based business cycle measures, we construct an approximation to Bayesian posterior model probabilities based on the Schwarz information criterion. The resulting model-averaged business cycle measure displays an asymmetric shape across NBER-dated recession and expansion phases. In particular, the business cycle measure has relatively small amplitude during mature expansions and substantial variation during and immediately following recessions.

The results for the model-averaged business cycle measure suggest a strong link between the output-gap notion of the business cycle and the NBER's alternating-phases notion. Specifically, NBER recessions are periods of significant transitory variation in output, while output in NBER expansions is dominated by movements in trend. This has potential relevance as a stylized fact to guide theoretical models of the business cycle. Beyond this link, we also find that the model-averaged business cycle measure is closely related to other measures of macroeconomic slack such as the unemployment rate and capacity utilization, even though the cycle is estimated by univariate analysis of real GDP. Taking these results together, we argue that the model-averaged business cycle measure captures a meaningful macroeconomic phenomenon and sheds more light on the nature of fluctuations in aggregate economic activity

Received for publication March 2, 2009. Revision accepted for publication September 20, 2010.

* Morley: University of New South Wales; Piger: University of Oregon.

We thank the editor, two anonymous referees, Marcellus Andrews, Frédérique Bec, Todd Clark, Drew Creal, Cees Diks, Neal Ghosh, Don Harding, Jinill Kim, Jim Nason, Zohra Rabah, Tara Sinclair, and Hermann van Dijk for helpful comments and suggestions. We also thank conference and seminar participants at the 2008 SNDE meeting, 2008 SEA meeting, 2009 NBER/NSF meeting, 2009 UC Riverside Conference on Business Cycles, 2009 WEA meeting, 2009 SCE meeting, 2009 Brookings Institution Conference on "Long Term Impacts of Short Term Fluctuations," Australian National University, BIS, Erasmus University Rotterdam, Ghent University, George Washington University, IMF Institute, Kansas City Fed, Tinbergen Institute Amsterdam, University of Melbourne, University of New South Wales, University of Sydney, and McGill University for helpful comments. We thank Jim Hamilton for providing code. J.M. acknowledges support from the Weidenbaum Center on the Economy, Government, and Public Policy at Washington University in St. Louis. The usual disclaimers apply.

than simply looking at the level or growth rates of real GDP.

The rest of this paper is organized as follows. Section II discusses different possible definitions of the business cycle, focusing on the output-gap view taken in this paper. Section III presents details on the model-based trend and cycle decomposition methods employed in our analysis, including the approach developed in Morley and Piger (2008). Section IV lays out the competing time series models of postwar U.S. real GDP, presents the implied business cycle measure for each model, and discriminates among these measures using model comparison based on information criteria. Section V reports hypothesis test results for a leading linear model against key nonlinear alternatives. Section VI presents the model-averaged measure of the business cycle and compares it to other measures of economic slack. Section VII concludes.

II. Definitions of the Business Cycle

In macroeconomics, fluctuations in economic activity are typically classified into three categories: long-run growth, the business cycle, and seasonal patterns. These different types of fluctuations may in fact be related to each other, but it can be useful to make some distinction among them. In this paper, we follow standard practice by considering seasonally adjusted data. This implicitly treats the seasonal patterns as independent or, at least, not marginally relevant for making inferences about long-run growth or business cycles, although we note the existence of an interesting literature on the influence of seasonal fluctuations on business cycles (see, for example, Wen, 2002).

So what is the business cycle as distinct from long-run growth? One notion put forth by the NBER is that the business cycle corresponds to an alternation between persistent phases of expansion and recession in economic activity that occur despite the positive average growth of economic activity in most industrialized countries. We refer to this notion of the business cycle as the alternating-phases definition.[1] One problem with this notion is that it is far from universal. Some countries have experienced many consecutive years of positive growth in the level of economic activity and thus have no business cycles in the strict NBER sense (for example, Japan in the early postwar period).

A more general notion of the business cycle is that it corresponds to all short-run fluctuations in economic activity (again, beyond seasonal movements), without a distinction made between whether they correspond to an increase or outright decline in activity. The problem with this definition is that it merely labels the analysis of higher-frequency variation in economic activity as "business cycle analysis," without saying whether there is anything meaningful about the business cycle as a macroeconomic phenomenon. For example, this notion begs the question of why any attention is paid to whether the NBER deems there to be a recession.

A third notion of the business cycle is that it represents the transitory fluctuations of the economy away from a long-run, or trend, level. In this paper, we argue that this output-gap definition provides the most useful notion of the business cycle.[2] It implies a construct—the transitory component of real economic activity—that can be measured for any economy and is potentially useful for forecasting, policymaking, and theory. We emphasize that nothing about this notion of the business cycle implies it is independent of long-run growth. Transitory fluctuations could be due to the same factors that drive long-run growth or due to independent factors. It is ultimately an empirical question as to how important these different underlying factors are. However, the key point is that the business cycle as the transitory component of economic activity is potentially an important macroeconomic phenomenon in its own right.

Before discussing methods of measuring the transitory component of economic activity in the next section, it is worth providing a more formal discussion of the output-gap view taken in this paper. First, following much of the literature, we use natural logarithms of U.S. quarterly real GDP, denoted y_t, as a measure of overall economic activity. We acknowledge that this measure has its limitations and does not always match up with the NBER's implicit measure of economic activity, but it does a reasonably good job on this front (see, for example, Harding & Pagan, 2002, and the subsequent literature on business cycle dating with real GDP). Then, given y_t, the output-gap definition of the business cycle corresponds to the idea that economic activity can be meaningfully decomposed into a trend and a cycle as follows:

$$y_t = \tau_t + c_t, \tag{1}$$

$$\tau_t = \tau_{t-1} + \eta_t^*, \tag{2}$$

$$c_t = \sum_{j=0}^{\infty} \psi_j \omega_{t-j}^*, \tag{3}$$

[1] Harding and Pagan (2005) also consider definitions of the business cycle that are closely related to those discussed here, albeit with somewhat different terminology.

[2] It is possible to make a further distinction between transitory movements at different frequencies. This is the approach taken when a spectral filter is applied to a time series with the goal of isolating fluctuations at, say, the one- to five-year horizon. However, it is important to note that these spectral filters are based on the assumption that the time series being analyzed follows a stationary process. When applied to integrated processes, part of the filter (a differencing operator) is used up transforming the series to something that could be thought of as stationary, leaving the rest of the filter to amplify fluctuations in the transformed series at different frequencies than originally intended (see Cogley & Nason, 1995, and Murray, 2003, on this point). While we acknowledge that the isolation of transitory movements at different frequencies is an interesting issue in business cycle analysis, we consider the initial isolation of transitory movements for measures of real economic activity as the more important and challenging task for macroeconomists given the basic premise that not all fluctuations are transitory.

where $\psi_0 = 1$, $\eta_t^* = \mu + \eta_t$, and $\omega_t^* = \bar{\omega} + \omega_t$, with η_t and ω_t following martingale difference sequences. The trend, τ_t, is the permanent component of y_t in the sense that the effects of the realized trend innovations, η_t^*, on the level of the time series are not expected to be reversed. By contrast, the cycle, c_t, is the transitory component of y_t in the sense that the Wold coefficients, ψ_j, are assumed to be absolutely summable such that the realized cycle innovations, ω_t^*, have finite memory. The parameter μ allows nonzero drift in the trend, while the parameter $\bar{\omega}$ allows a nonzero mean in the cycle, although the mean of the cycle is not identified from the behavior of the time series alone, as different values for $\bar{\omega}$ all imply the same reduced-form dynamics for Δy_t, with the standard identification assumption being that $\bar{\omega} = 0$.

Whether the permanent and transitory components in equation (1) are meaningful macroeconomic phenomena is ultimately an empirical question, although it is clear that the trend should embody the steady-state effects of the factors that drive long-run growth in economic activity.[3] Such factors might also have transitory effects, so we do not assume the permanent and transitory shocks are uncorrelated (see Morley, Nelson, & Zivot, 2003, on this point). Meanwhile, the business cycle may be related to other macroeconomic phenomena such as inflation. However, to allow empirical tests of such relationships, we do not assume them a priori.

In this paper, we consider the setting where some of the parameters describing the process in equations (1) to (3) are regime switching, as discussed in Morley and Piger (2008), and where some of the parameters undergo structural breaks. In addition to finite-order unobserved-components (UC) models of the process in equations (1) to (3), we also consider processes for which there is no finite-order autoregressive moving-average (ARMA) representation of the Wold form in equation (3). Specifically, we conduct trend and cycle decomposition based on reduced-form forecasting models that capture the autocovariance structure for a general process as in equations (1) to (3), regardless of whether the process has a finite-order UC representation. Given a forecasting model that captures the autocovariance structure of a process as in equations (1) to (3), the methods we employ provide optimal estimates (in a minimum mean-squared-error sense) of trend and cycle.

III. Trend/Cycle Decomposition Based on Time Series Models

A. The Beveridge-Nelson Decomposition

There are many different approaches to trend and cycle decomposition. In terms of the output-gap definition of the business cycle as transitory deviations away from trend, a particularly useful and general approach is the Beveridge-Nelson (BN) decomposition. The BN measure of trend is

$$\hat{\tau}_t^{BN} \equiv \lim_{j\to\infty}\left\{E^M\left[y_{t+j}|\Omega_t\right] - j \cdot E^M[\Delta y_t]\right\}, \qquad (4)$$

where $E^M[\cdot]$ is the expectations operator with respect to a forecasting model and Ω_t is the set of relevant and available information observed up to time t. In words, the BN trend is the long-horizon conditional forecast of the time series minus any deterministic drift. The intuition for the BN measure of trend is that as the forecasting horizon extends to infinity, a long-horizon forecast of a time series should no longer be influenced by the transitory component that exists at time t, and therefore should reflect only the trend component.

Both the conditional and unconditional expectations in equation (4) are usually straightforward to calculate (either analytically or by simulation) given a forecasting model. Morley (2002) and Clarida and Taylor (2003) provide discussion and examples, while appendix A provides the relevant formulas for the class of forecasting models to which we apply the BN decomposition in the next section. Meanwhile, the BN trend provides an optimal estimate of the underlying trend of an integrated process in the following circumstances. First, the time series under analysis conforms to the trend and cycle process in equations (1) to (3), with constant drift, μ, constant Wold coefficients, ψ_j, and the mean of the cycle innovations equal to 0, $\bar{\omega} = 0$. Second, the forecasting model captures the autocovariance structure of the process such that $E^M\left[y_{t+j}|\Omega_t\right] = E\left[y_{t+j}|\Omega_t\right]$.

This second requirement highlights the fact that accurate measurement of the business cycle requires an accurate forecasting model. This is important because it justifies our choice not to limit our consideration only to finite-order UC models in order to capture the process in equations (1) to (3). Such models represent a mere subset of all possible time series models and can place binding restrictions on the autocovariance structure of a given time series process. By considering a broader set of models, we aim to get $E^M[y_{t+j}|\Omega_t]$ as close as possible to $E[y_{t+j}|\Omega_t]$.

B. Regime-Dependent Steady-State Approach

As shown in Morley and Piger (2008), the BN trend does not generally provide an optimal or even unbiased estimate of trend when the process in equations (1) to (3) has regime-switching parameters, even if the forecasting model is correctly specified such that $E^M[y_{t+j}|\Omega_t] = E[y_{t+j}|\Omega_t]$. As an alternative, we consider the regime-dependent steady-state (RDSS) approach from Morley and Piger (2008) that generalizes the BN decomposition to provide optimal estimates when the underlying trend or cycle is regime switching.

The RDSS approach involves constructing long-horizon forecasts conditional on sequences of regimes and then

[3] To be clear with our terminology, by "steady state" we have in mind the level to which the process would gravitate in the absence of future permanent or transitory innovations.

marginalizing over the distribution of the unknown regimes. Specifically, the RDSS measure of trend is

$$\hat{\tau}_t^{RDSS} \equiv \sum_{\tilde{S}_t} \left\{ \hat{\tau}_t^{RDSS}\left(\tilde{S}_t\right) \cdot p^M\left(\tilde{S}_t | \Omega_t\right) \right\}, \quad (5)$$

$$\hat{\tau}_t^{RDSS}\left(\tilde{S}_t\right) \equiv \lim_{j \to \infty} \left\{ E^M\left[y_{t+j} \middle| \{S_{t+k} = i^*\}_{k=1}^{j}, \tilde{S}_t, \Omega_t \right] - j \cdot E^M\left[\Delta y_t \middle| \{S_t = i^*\}_{-\infty}^{\infty} \right] \right\}, \quad (6)$$

where $\tilde{S}_t = \{S_t, ..., S_{t-m}\}'$ is a vector of relevant current and past regimes for forecasting a time series, $p^M(\cdot)$ is the probability distribution with respect to the forecasting model, S_t is an unobserved Markov state variable that takes on N discrete values according to a fixed transition matrix, and i^* is the "normal" regime in which the mean of the transitory component is assumed to be 0. The choice of "normal" regime i^* is necessary for identification. However, unlike the BN decomposition, there is no implicit assumption that the cycle is unconditionally mean 0. Meanwhile, for a given forecasting model, the probability weights in equation (5), $p^M\left(\tilde{S}_t | \Omega_t\right)$, can be obtained from the filter given in Hamilton (1989). Appendix A provides the relevant formulas for constructing the expectations in equation (6) for the regime-switching models considered in this paper.

As long as $E^M[y_{t+j}|\Omega_t] = E[y_{t+j}|\Omega_t]$, the RDSS approach will provide an optimal estimate of trend when the process in equations (1) to (3) has regime-switching parameters (see Morley & Piger, 2008, for full details of the RDSS approach). Meanwhile, the RDSS approach is general in the sense that it simplifies to the BN decomposition in the absence of regime switching.

C. UC Models

A direct way to conduct trend and cycle decomposition is to consider a finite-order parametric specification for the Wold form of the transitory component in equation (3). For example, a standard assumption is a finite-order stationary AR process,

$$\phi(L)c_t = \omega_{t-j}^*, \quad (3')$$

where $\phi(L)$ denotes a lag polynomial with roots outside the unit circle and, again, the standard identification assumption for the mean of the cycle is that $\bar{\omega} = 0$. Assuming that the shocks to the trend and the cycle in equations (2) and (3′) are Gaussian $(\eta_t, \omega_t)' \sim N(0, \Sigma_{\eta\omega})$), the Kalman filter can be employed to make optimal inferences about the trend and cycle. As Morley et al. (2003) discussed, the inferences based on the Kalman filter will be the same as those based on the BN decomposition given equivalent models of the autocovariance structure of a time series.

It is possible to extend linear UC models to allow for regime-switching parameters. For example, Lam (1990) considers the case where the drift parameter is regime switching ($\mu = \mu(S_t)$ in equation [2]). Kim and Nelson (1999a) consider the case where the mean of the innovations to the cycle is regime switching—$\bar{\omega} = \bar{\omega}(S_t)$ in equation (3′). As with the RDSS approach, it is necessary for identification to assume a "normal" regime i^* in which the cycle is mean 0. Meanwhile, the Kalman filter is no longer available for regime-switching UC models, but estimates of trend and cycle can still be obtained by using a Bayesian posterior simulator such as the Gibbs sampler or Kim's (1994) analytical approximation to the optimal filter. Finally, it is worth noting that the RDSS approach and inference based on optimal filtering will be the same given equivalent models of the autocovariance structure of the time series, while the BN decomposition in general will be different and not optimal.

IV. Model-Based Measures of the U.S. Business Cycle

A. Models of U.S. Real GDP

To keep the scope of our analysis manageable, we consider only univariate models of postwar U.S. real GDP. We justify this focus on univariate models in part because trend and cycle decomposition is usually considered as a prior step to cross-series analysis. For example, researchers are often interested in whether and how trend and cycle components of one time series are related to the trend and cycle components of many other series. As a general method, then, it is particularly useful if trend and cycle decomposition can be applied first at a univariate level and then the resulting measures of trend and cycle considered in different multivariate settings. For instance, this is the approach taken in studies that use the Hodrick-Prescott filter or a bandpass filter and could help explain their popularity. Meanwhile, it should be noted that the univariate models considered here capture a wide range of possibilities about the predictability of postwar U.S. real GDP.

In terms of linear models, we consider AR models for the first differences of y_t:

$$\phi(L)(\Delta y_t - \mu) = e_t, \quad (7)$$

where $\phi(L)$ is pth order with p set to certain values ranging from 0 to 12. We consider versions of the AR models with Gaussian errors ($e_t \sim N(0, \sigma_e^2)$) or Student t errors ($e_t \sim t(\nu, 0, \sigma_e^2)$). Beyond the AR models, we also consider three UC models. The first model (UC-HP) is due to Harvey and Jaeger (1993) and corresponds to the Hodrick-Prescott filter with a smoothing parameter of 1,600.[4] The second model (UC-0) has a standard UC specification as in equations (1), (2), and (3′), with an independent AR(2) cycle ($\eta_t \sim N(0, \sigma_\eta^2)$, $\omega_t \sim N(0, \sigma_\omega^2)$, and $E[\eta_t, \omega_t] = 0$). The third model (UC-UR)

[4] Harvey and Jaeger's (1993) UC model assumes a random walk with drift for the permanent component, where the drift itself follows a random walk, plus noise for the transitory component (an AR(0) cycle). The variance of the shock to the permanent component is assumed to be 0, while the variance of the shock to the drift is assumed to be 1/1600 times as large as the variance of the noise, which is freely estimated.

has the same structure as the second model except that, following Morley et al. (2003), it allows correlation between permanent and transitory movements by assuming a general variance-covariance matrix $\Sigma_{\eta\omega}$ for the shocks. Because the Kalman filter assumes Gaussian shocks, we do not consider Student *t* errors for the UC models.

In terms of nonlinear models, we consider Hamilton's (1989) Markov-switching model and different versions of Kim et al's. (2005) "bounceback model." If the Hamilton model can be said to correspond to L-shaped recessions, with the economy growing from a permanently lower level following the end of a recession, the bounceback models allow a postrecession recovery phase (see Sichel, 1994), with the three cases of U-shaped recessions, V-shaped recessions, and recoveries that are proportional to the "depth" of the preceding recession. Each of these nonlinear models can be expressed as an AR model with a regime-switching mean that potentially depends on the current and *m* lagged states:

$$\phi(L)(\Delta y_t - \mu_t) = e_t, \tag{8}$$

$$\mu_t = \mu(S_t, ..., S_{t-m}), \tag{9}$$

where $S_t = \{0,1\}$ is a Markov state variable with fixed continuation probabilities $\Pr[S_t = 0|S_{t-1} = 0] = p_{00}$ and $\Pr[S_t = 0|S_{t-1} = 0] = p_{11}$. The Hamilton and bounceback models differ by their specifications for the time-varying mean:

- Hamilton (H):

$$\mu_t = \gamma_0 + \gamma_1 S_t \tag{10}$$

- U-shaped recessions (BBU):

$$\mu_t = \gamma_0 + \gamma_1 S_t + \lambda \sum_{j=1}^{m} \gamma_1 S_{t-j} \tag{11}$$

- V-shaped recessions (BBV):

$$\mu_t = \gamma_0 + \gamma_1 S_t + (1 - S_t)\lambda \sum_{j=1}^{m} \gamma_1 S_{t-j} \tag{12}$$

- Recovery based on depth (BBD):

$$\mu_t = \gamma_0 + \gamma_1 S_t + \lambda \sum_{j=1}^{m} (\gamma_1 + \Delta y_{t-j}) S_{t-j} \tag{13}$$

where the state $S_t = 1$ is labeled as the low-growth regime by assuming $\gamma_1 < 0$. Following Kim et al. (2005), we assume $m = 6$ for the bounceback models, which allows recoveries to persist for up to six quarters following the end of a recession. Again, for these nonlinear AR models, we consider cases of both Gaussian errors and Student *t* errors.

In terms of nonlinear UC models, we consider Kim and Nelson's (1999a) version of Milton Friedman's "plucking" model (UC-FP-0) and a version due to Sinclair (2009) that allows correlation between permanent and transitory shocks (UC-FP-UR). These models augment the linear UC-0 and UC-UR models described earlier by allowing a regime-switching mean of the cyclical component in equation (3′),

$$\bar{\omega} = \tau S_t, \tag{14}$$

where S_t is defined as before for the nonlinear AR models and the state $S_t = 1$ is labeled by assuming $\tau < 0$. As in the linear case, we consider Gaussian shocks only for the nonlinear UC models.

While there are many other nonlinear models, these cover the range of possibilities in terms of whether recessions are permanent or transitory. The Hamilton model assumes the effects of regime switches into recessions are completely permanent, the plucking model assumes they are completely transitory, and the bounceback models allow both possibilities and everything in between.

For both linear and nonlinear time series models of postwar U.S. real GDP, a vast literature documents evidence of structural breaks. In particular, the evidence for a structural break in volatility sometime during 1984 (the so-called Great Moderation) is as close to incontrovertible as it gets in time series analysis of macroeconomic data, and several studies have pointed out the importance of accounting for this volatility change when estimating regime-switching models (Kim and Nelson, 1999b; McConnell & Perez-Quiros, 2000). Although less overwhelming than the evidence for the Great Moderation, there is some evidence for a reduction in mean growth rates in the early 1970s (the so-called productivity slowdown) that has been considered in a number of studies (Perron, 1989; Bai, Lumsdaine, & Stock, 1998). In a recent paper, Perron and Wada (2009) argue that controlling for the productivity slowdown is crucially important for U.S. business cycle measurement. They show that measures of the business cycle for different UC models are less sensitive to model specification once a break in the long-run average growth rate of U.S. real GDP is allowed in 1973. Thus, for all models under consideration, we allow a break in long-run growth in the first quarter of 1973 and a break in volatility in the second quarter of 1984. To keep the addition of parameters across models the same, we accommodate each structural break by a single parameter.[5] Table 1 details how we parameterize the structural breaks for each model.

For the trend and cycle decomposition given the linear models, we use the BN decomposition or, in the case of the UC models, the Kalman filter. Note again that the filtered inferences from the Kalman filter are equivalent to the BN decomposition using the corresponding reduced form of the UC model. For trend and cycle decomposition given the nonlinear forecasting models, we use the RDSS approach

[5] More complicated patterns of structural change yielded only small improvements in fit and had little effect on inferences about the business cycle. Meanwhile, for the set of preferred models that emerge from the model comparison, the business cycle estimates were robust to the exclusion of one or both breaks. However, according to the information criteria, specifications excluding both structural breaks were strongly dominated by those including breaks for all models under consideration. Thus, we focus on the results for models with structural breaks.

TABLE 1.—PARAMETERIZATION OF STRUCTURAL BREAKS

Models	Average Growth Break	Volatility Break
AR(p)	$\mu^{\text{post-1973:1}} = \delta_1 \mu^{\text{pre-1973:1}}$	$\sigma_e^{\text{post-1984:2}} = \delta_2 \sigma_e^{\text{pre-1984:2}}$
UC-0, UC-UR, UC-FP-0, UC-FP-UR	$\mu^{\text{post-1973:1}} = \delta_1 \mu^{\text{pre-1973:1}}$	$\Sigma_{\eta\omega}^{\text{post-1984:2}} = \delta_2 \Sigma_{\eta\omega}^{\text{pre-1984:2}}$
UC-HP	NA[a]	$\Sigma_{\eta\omega}^{\text{post-1984:2}} = \delta_2 \Sigma_{\eta\omega}^{\text{pre-1984:2}}$
H, BB	$\gamma_0^{\text{post-1973:1}} = \delta_1 \gamma_0^{\text{pre-1973:1}}$	$\gamma_1^{\text{post-1984:2}} = \delta_2 \gamma_1^{\text{pre-1984:2}}$ $\sigma_e^{\text{post-1984:2}} = \delta_2 \sigma_e^{\text{pre-1984:2}}$

The "Average Growth Break" column details how a one-time break in long-run average growth in the first quarter of 1973 is incorporated in each of the models, and the "Volatility Break" column shows this information for a one-time break in volatility parameters in the second quarter of 1984. Each break is parameterized such that it adds a single free parameter to each model. Following Kim and Nelson (1999b), a break in the difference in growth rates across regimes for the Hamilton and Bounceback models is included as part of the volatility break. [a]The drift parameter in the UC-HP model follows a random walk process. Because this allows time variation in long-run average growth, we do not allow for additional structural breaks in the drift parameter.

TABLE 2.—INFORMATION CRITERIA FOR LINEAR MODELS

Model	Log Likelihood	Number of Parameters	Akaike Information Criterion	Schwarz Information Criterion	Posterior Model Probability
AR(0)	−297.50	4	−301.50	−308.46	0.00
AR(1)	−287.00	5	−292.00	−300.70	0.17
AR(2)	−283.97	6	−289.97	**−300.40**	0.23
AR(4)	−281.52	8	−289.52	−303.42	0.01
AR(8)	−280.92	12	−292.92	−313.78	0.00
AR(12)	−273.66	16	−289.66	−317.47	0.00
AR(0)-t	−296.64	5	−301.64	−310.33	0.00
AR(1)-t	−286.29	6	−292.29	−302.72	0.02
AR(2)-t	−283.51	7	−290.51	−302.68	0.02
AR(4)-t	−280.94	9	−289.94	−305.58	0.00
AR(8)-t	−280.37	13	−293.37	−315.96	0.00
AR(12)-t	−273.62	17	−290.62	−320.17	0.00
UC-HP	−505.15	2	−507.15	−510.63	0.00
UC-0	−282.00	7	−289.00	−301.17	0.11
UC-UR	−280.70	8	**−288.70**	−302.60	0.03

Maximum likelihood estimation is based on the conditional likelihood for 1947:Q2-2006:Q4. Observations prior to 1947:Q2 are backcast based on the mean growth rate. The AIC and SIC are formulated such that the highest value (in bold) represents the preferred model. Posterior model probabilities are based on the asymptotic approximation given by the SIC, as discussed in section VI.

or, in the case of the nonlinear UC models, the Kim (1994) filter, which combines the Kalman filter with Hamilton's (1989) filter for Markov-switching models. For the nonlinear models, we follow Kim and Nelson (1999a) and Sinclair (2009) by assuming $i^* = 0$, which corresponds to an assumption that the cycle is mean 0 in expansions.[6]

B. Estimates and Comparison of Business Cycle Measures

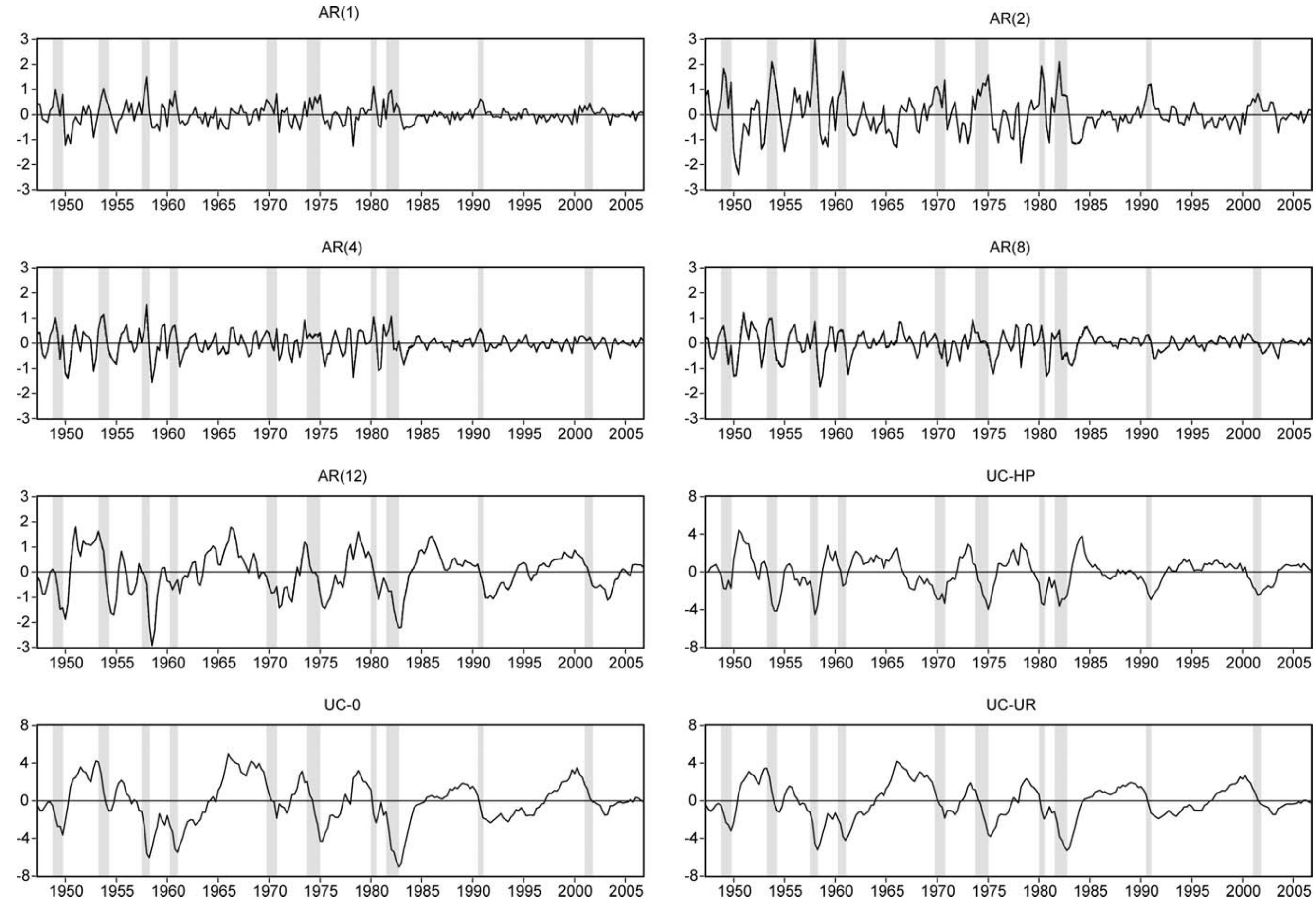
The raw data are seasonally adjusted quarterly U.S. real GDP for the sample period of 1947:Q1 to 2006:Q4 and were taken from the St. Louis Fed (FRED) database. We conduct maximum likelihood estimation (MLE) for all of the models and use the Akaike and Schwarz information criteria (AIC and SIC) for model comparison.[7] To facilitate model comparison, we need to ensure that the adjusted sample period is equivalent for all of the models. Complicating matters is the practical difficulty of conducting exact MLE for the nonlinear AR models. Our solution to this problem is to backcast a suitable number of observations prior to 1947:Q1 based on the long-run average growth rate. We then conduct conditional MLE based on the same adjusted sample of 1947:Q2-2006:Q4 for 100 times the first differences of the natural logs of real GDP.

Table 2 reports AIC and SIC results for the various linear models, and figure 1 reports the corresponding cycle measures for all but the AR(0) model, which has no cycle by assumption. Beginning with the results for the AR models, AIC picks the AR(4) model with Gaussian errors, while SIC picks the AR(2) model with Gaussian errors. Looking at figure 1, we can see that the AR(2) and AR(4) cycles are small, noisy, and typically positive during NBER recessions. The reason for this counterintuitive result is that both models imply positive serial correlation at short horizons. Specifically, when output falls in a recession, there is a prediction of further declines (or at least below-average growth) in the short run, suggesting that output is above its long-run (trend) level. By contrast, the AR(12) model produces a more traditional-looking cycle that typically turns negative during NBER-dated recessions. In this case, the model implies negative serial correlation at longer horizons. Thus, when output falls in a recession, there is a prediction of compensating above-average growth at some point in the future, suggesting that output is below its long-run level. The very different cycle for the AR(12) model is notable in

[6] It should be noted that this assumption places no a priori restriction on the unconditional mean of the cycle for the bounceback models and does not affect the general shape of the cycle. Also, the results for U.S. real GDP suggest that the assumption of a mean 0 cycle in expansions is the only assumption consistent with the steady-state notion that output is at its trend level when the change in the cycle (which depends on the shape, but not the level of the cycle) remains at 0 for an extended period of time.

[7] We define AIC and SIC as in Davidson and MacKinnon (2004). Specifically, goodness of fit is measured by the log likelihood, and the offsetting penalty for AIC is the number of parameters, while the penalty for SIC is the number of parameters times one-half of the log of the sample size (natural logarithms in both cases). This definition is such that larger values of the information criteria are preferred.

FIGURE 1.—MEASURES OF THE U.S. BUSINESS CYCLE BASED ON LINEAR MODELS



NBER recessions are shaded.

part because even though the model is heavily discounted by SIC, it has a sizable improvement in likelihood over even the AR(8) model and is reasonably close to the low-order AR models when considering AIC.

In terms of the UC models, it is interesting to note how similar the HP cycle looks to the AR(12) cycle. However, there are very different scales for these two cycles, with the HP cycle being much larger in amplitude. Furthermore, the UC-HP model has an extremely poor fit as judged by either of the information criteria, suggesting that the autocovariance structure implied by the UC-HP model is strongly at odds with the data. In terms of model comparison, SIC favors the UC-0 model, while AIC slightly favors the UC-UR model. Interestingly, from figure 1, both the UC-0 and UC-UR cycles are large and persistent and have a similar pattern to the AR(12) cycle.

Comparing across all the linear models, the model selection criteria produce a mixed signal about the nature of the business cycle. The preferred model, as judged by AIC, is the UC-UR model, which, from figure 1, produces a large, traditional business cycle that implies an important role for transitory fluctuations. The preferred model, as judged by SIC, is the AR(2) model with Gaussian errors. In contrast to the UC-UR model, the AR(2) produces a small, nontraditional business cycle, implying that most short-run fluctuations in output are permanent.

Before turning to the nonlinear models, it is worth comparing the results obtained for the linear models to the conclusions of Perron and Wada (2009), who argue that the sensitivity of model-based measures of the business cycle is due to a failure to account for structural breaks. In particular, they show that the cycles implied by different methods and specifications of linear models look more similar once a one-time break in the long-run growth rate of U.S. real GDP is allowed in 1973. While the similarity in the implied cycles generated by the UC-0 and UC-UR models corroborates the Perron and Wada finding, the discrepancy between the implied cycles generated by the low-order AR models and the UC models does not. Thus, allowing for structural breaks does not in fact resolve the sensitivity of business cycle measures to model specification.

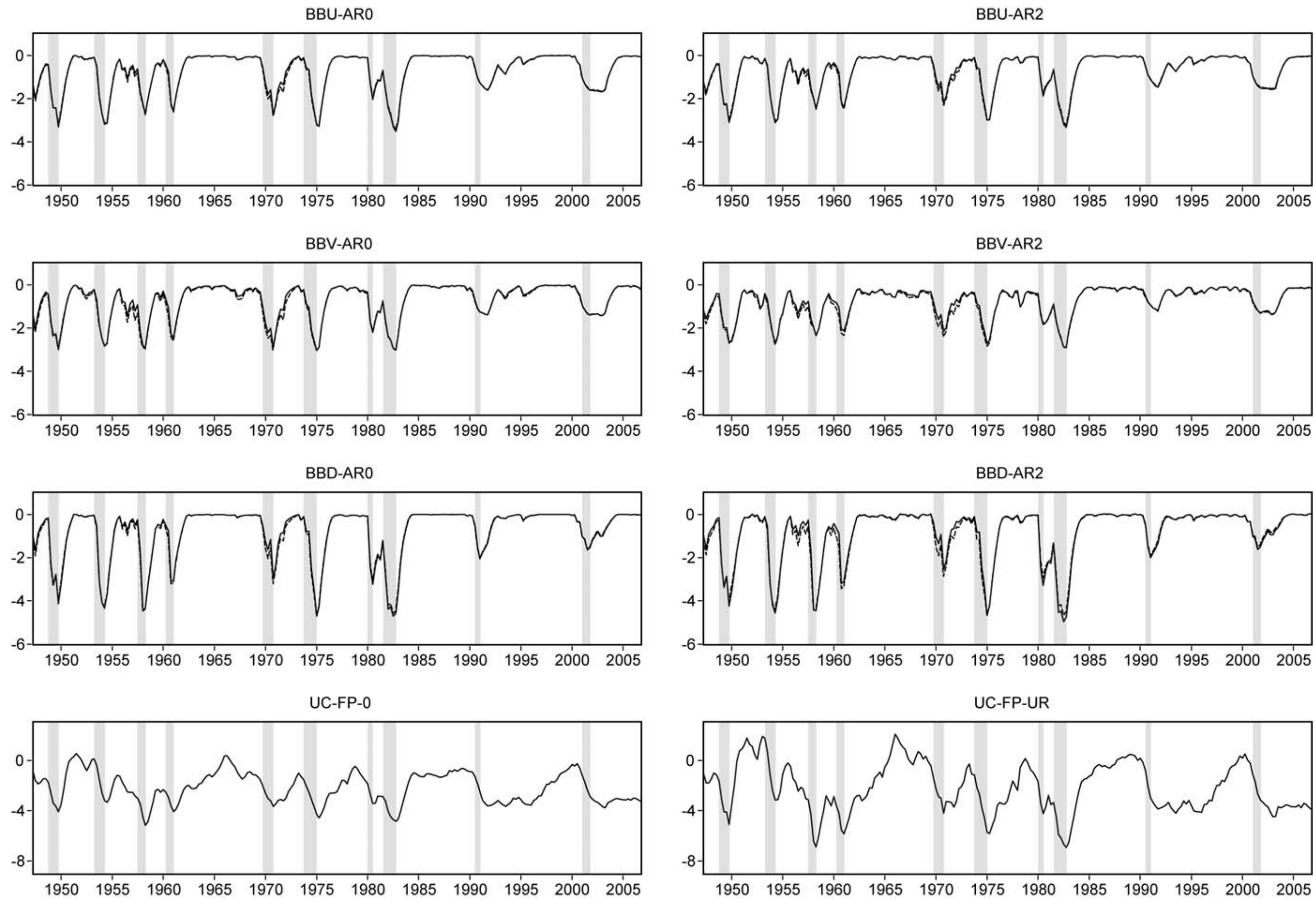
Table 3 reports AIC and SIC results for the various nonlinear models, and figure 2 reports the corresponding cycle measures for all but the Hamilton model, which has only a linear cycle by assumption. Beginning with the nonlinear AR models, the results in table 3 suggest that the bounce-back model, in all of its versions, is strongly preferred to the corresponding Hamilton model. Evidently the idea that

TABLE 3.—INFORMATION CRITERIA FOR NONLINEAR MODELS

Model	Log Likelihood	Number of Parameters	Akaike Information Criterion	Schwarz Information Criterion	Posterior Model Probability
H-AR0	−286.13	7	−293.13	−305.29	0.00
H-AR2	−282.71	9	−291.71	−307.35	0.00
BBU-AR0	−280.34	8	−288.34	−302.25	0.03
BBU-AR2	−279.77	10	−289.77	−307.15	0.00
BBV-AR0	−281.24	8	−289.24	−303.15	0.01
BBV-AR2	−279.33	10	−289.33	−306.71	0.00
BBD-AR0	−277.95	8	**−285.95**	**−299.86**	0.33
BBD-AR2	−277.14	10	−287.14	−304.53	0.00
H-AR0-t	−285.83	8	−293.83	−307.74	0.00
H-AR2-t	−282.27	10	−292.27	−309.65	0.00
BBU-AR0-t	−280.19	9	−289.19	−304.84	0.00
BBU-AR2-t	−279.62	11	−290.62	−309.75	0.00
BBV-AR0-t	−280.88	9	−289.88	−305.53	0.00
BBV-AR2-t	−278.97	11	−289.97	−309.09	0.00
BBD-AR0-t	−277.59	9	−286.59	−302.23	0.03
BBD-AR2-t	−276.53	11	−287.53	−306.65	0.00
UC-FP-0	−281.38	10	−291.38	−308.76	0.00
UC-FP-UR	−280.09	11	−291.09	−310.21	0.00

Maximum likelihood estimation is based on the conditional likelihood for 1947:Q2-2006:Q4. Observations prior to 1947:Q2 are backcast based on the mean growth rate. The AIC and SIC are formulated such that the highest value (in bold) represents the preferred model. Posterior model probabilities are based on the asymptotic approximation given by the SIC, as discussed in section VI.

FIGURE 2.—MEASURES OF THE U.S. BUSINESS CYCLE BASED ON NONLINEAR MODELS



NBER recessions are shaded.

the regime switches correspond to only permanent movements in the level of output is not supported by the data. The nonlinear UC models fare somewhat better than the Hamilton model, at least according to AIC. However, the various versions of the bounceback model are generally preferred to the nonlinear UC models.

Comparing across all of the nonlinear models, both AIC and SIC choose the depth version of the bounceback model (BBD-AR0) with no linear dynamics and Gaussian errors. However, from the perspective of measuring the business cycle, all versions of the bounceback model yield similar results. From figure 2, the different bounceback models all imply cycles that have a similar shape and display a distinct asymmetry in the form of "deepness" (see Sichel, 1993) across NBER-dated recession and expansion phases. Specifically, the implied cycles from the bounceback models display little or no variation during mature expansions, which suggests that most variation in output during expansions is due to fluctuations in trend. However, in recessions and their immediate aftermath, the bounceback model cycles display considerable variation. Taken together, this implies an asymmetry in the extent of transitory fluctuations in output across expansion and recession phases. Furthermore, this asymmetry suggests a direct link between the NBER definition of the business cycle and the transitory component of U.S. real GDP. In terms of recessions, the NBER appears to be identifying periods in which there are substantial negative transitory fluctuations in real economic activity.

Finally, we turn to the comparison between linear and nonlinear models. According to both AIC and SIC, the preferred model is the BBD-AR0 model with Gaussian errors.[8] The AIC results suggest that the bounceback models are dominant, as the only models with an AIC statistic close to that for the BBD-AR0 are other bounceback models. Thus, if we restrict attention to AIC and given the robustness of the business cycle measures to alternative bounceback models, the business cycle is adequately measured using the BBD-AR0 model. However, the SIC results suggest a greater degree of model uncertainty, with several models that are relatively close competitors to the bounceback models. The final columns of tables 2 and 3 present posterior model probabilities implied by SIC, the construction of which will be discussed in section VI. These posterior probabilities demonstrate that while the BBD-AR0 model receives the highest posterior probability (33%) of any individual model, several other models receive nonnegligible posterior probability, including the linear AR(1) model (17%), linear AR(2) model (23%), and linear UC-0 model (11%). If we aggregate probabilities across groups of models that produce similar business cycle measures, essentially all posterior probability is accounted for by three groups: bounceback models (40%), low-order AR models (44%), and linear UC models (14%). In terms of measuring the business cycle, this model uncertainty is relevant because these three groups of models yield starkly different estimates of the business cycle. In particular, while the bounceback models give an estimate of the cycle that is large and asymmetric, low-order AR models yield small symmetric cycles and linear UC models yield large symmetric cycles. Thus, while the model comparison was successful at compressing the set of relevant models, we are still left with a fair degree of uncertainty regarding the appropriate measure of the business cycle.

V. Tests of Nonlinearity

The standard information criteria choose a nonlinear model with no linear dynamics, namely, the BBD-AR0 version of the bounceback model, as the single best model. However, the support for nonlinearity is far from definitive because the information criteria suggest some linear models that are close competitors. Furthermore, as is evident in tables 2 and 3, when comparing a more general version of the bounceback model that incorporates AR(2) dynamics to its nested linear AR(2) counterpart, SIC favors the linear model (although AIC continues to favor the nonlinear model). Thus, it is not clear if the bounceback model would also be supported by formal statistical tests of nested models. In this section, we take up formal testing of nonlinearity within the context of a few of the key models considered in the previous section. In particular, we consider a null hypothesis of a linear AR(2) model and compare it to the nonlinear alternatives of the Hamilton model and the bounceback models with AR(2) dynamics.

Testing for nonlinearity of the Markov-switching form is difficult due to the presence of unidentified nuisance parameters under the null hypothesis of linearity and the singularity of the information matrix at the null. There have been different proposed tests to address this nonstandard environment, most notably by Hansen (1992) and Garcia (1998). Recently, Carrasco, Hu, and Ploberger (2007) developed a relatively straightforward information-matrix-based test (the CHP test hereafter) that is optimal for local alternatives to linearity and requires estimation only under the null hypothesis of linearity. While estimation under the null might seem like a small advantage given that we estimated the alternative models in the previous section, it is helpful because the test requires parametric bootstrap experiments to assess statistical significance. The bootstrap experiments are made easier by having to estimate models only under the null of linearity given data generated under the null. But, there are some limitations in terms of what alternatives can be considered with the CHP test. Thus, at the end of this section, we also discuss results for parametric bootstrap experiments to assess the significance of a likelihood ratio

[8] It is notable that the findings in favor of the bounceback model are robust to allowing for Student *t* errors, implying that it is the ability of the model to capture nonlinear dynamics, rather than fat tails in the unconditional distribution of output growth that explains its empirical success. We found more support for Student *t* errors when considering models without structural breaks. However, this directly suggests that it is really the structural breaks (especially the Great Moderation) that matter for output growth rather than fat tails in the error distribution. Meanwhile, the estimates for the nonlinear models imply persistent regimes, suggesting that it really is nonlinear dynamics rather than asymmetric shocks that are important. Indeed, while the likelihood-based analysis here does not directly consider asymmetric shocks, Morley, Piger, and Tien (2010) show that linear time series models with shocks based on empirical distributions are unable to reproduce key business cycle features that bounceback models with parametric and symmetric shocks are able to reproduce.

(LR) test for nonlinearity.[9] In this case, the bootstrap experiments require estimation under the alternative. We address difficulties in estimating under the alternative by considering a grid of possible values for the continuation probabilities (see Kim et al., 2005) and a large number of starting values for MLE.

The CHP test can be applied to a broad set of random coefficient models, the most prominent of which are models with Markov-switching parameters. In addition to presenting the general test, Carrasco et al. (2007) discuss how to implement the test in the specific case where parameters depend on only the current realization of a two-state Markov-switching process (also see Hamilton, 2005, for an accessible discussion of how to implement the CHP test in this case). In appendix B, we provide details on how to implement the CHP test for a broader range of Markov-switching models, including the Hamilton and bounceback models considered in this paper, for which parameters can depend on current and lagged values of the state variable.

For our tests, we consider a linear AR(2) model as the null hypothesis. As in the previous section, we assume Gaussian or Student *t* errors and allow structural breaks in mean and variance. In terms of alternatives, we consider Hamilton's (1989) model, which implies L-shaped recessions and the U-shape-recession version of the bounceback model (BBU). Note that we do not consider the V-shape-recession version of the bounceback model (BBV) because the CHP test requires that the regime-switching parameters be linear function of the state variables, while the BBV model has the mean depend in part on the product of the current and lagged states. The depth version of the bounceback model (BBD) is also difficult to cast into the CHP test framework due to the interaction between the lagged states and lagged growth rates. Thus, we consider a bootstrap LR test instead of the CHP test for the BBD alternative.

Table 4 presents the results for the tests of linearity. For the Hamilton model as an alternative, we are unable to reject linearity using the CHP test. However, for the BBU model as an alternative, the bootstrap *p*-value of the CHP test is 0.03, meaning that we can reject the null of linearity at the standard 5% level. Likewise, the bootstrap *p*-value is 0.03 for the LR test with the BBD model as an alternative. These results are robust to consideration of models with Student *t* errors. Thus, using formal hypothesis tests, there is evidence for nonlinearity given nonlinear alternatives that allow high-growth recoveries following the end of recessions.

Yet while we are able to reject linearity for specific nonlinear alternatives, it must be acknowledged that the more alternatives we consider, the more we are faced with a potential size distortion in our overall test of linearity. Thus, in terms of measuring the business cycle, we are left with the unsatisfactory situation that our inferences depend crucially on close to knife-edge test results about whether a linear model or nonlinear model provides a better description of the autocovariance structure of U.S. real GDP growth. Taken together with the model uncertainty demonstrated by the comparisons based on information criteria, an obvious response is to construct a model-averaged measure of the business cycle that weights alternative business cycle measures in a way that incorporates model uncertainty. This is the approach that we take in the next section.

[9] Di Sanzo (2009) uses Monte Carlo analysis to investigate the small sample properties of some statistical tests for Markov switching, including the CHP test and a parametric bootstrap LR test. He finds that both tests have good size properties, but that the LR test has higher power for the specific data-generating processes considered in his study.

TABLE 4.—TESTS OF NONLINEARITY

Null		Alternatives: L-Shape (Hamilton)	U-Shape (BBU)	Depth (BBD)
AR(2)	Test statistic	0.07	2.14	13.32
	(*p*-value)	(0.52)	(0.03)	(0.03)
	95% critical value	0.35	1.99	12.20
AR(2)-t	Test statistic	0.27	2.36	13.40
	(*p*-value)	(0.19)	(0.03)	(0.03)
	95% critical value	0.79	1.98	12.11

The test statistics for the L-shaped and U-shaped recession alternatives are based on Carrasco et al. (2007). The test statistics for the depth-based recovery alternative are likelihood ratio statistics based on estimation using a grid for the continuation probabilities. All *p*-values and critical values are based on parametric bootstrap experiments with 499 simulations.

VI. A Model-Averaged Measure of the Business Cycle

To construct a model-averaged measure of the business cycle, we take a Bayesian approach to model uncertainty and assign a posterior probability that each model is true. The Bayesian model-averaged measure of the business cycle, denoted $\tilde{c}_t$, is then a probability-weighted sum of the model-specific business cycle measures:

$$\tilde{c}_t = \sum_{i=1}^{N} c_{i,t} \Pr(M_i|y), \tag{15}$$

where i indexes the N models under consideration, $c_{i,t}$ is the business cycle measure for model i, M_i is an indicator for model i, and $\Pr(M_i|y)$ denotes the posterior probability that model i is true, conditional on the data, y.

From a Bayesian perspective, the model-averaged measure in equation (15) is the optimal solution to incorporating model uncertainty under certain conditions. Specifically, assuming the set of models under consideration is exhaustive, Min and Zellner (1993) show that $\tilde{c}_t$ minimizes expected predictive squared error loss. However, it is worth mentioning that even without taking a Bayesian viewpoint, constructing a model-averaged measure of the business cycle has some justification. In particular, the BN and RDSS methods that we use to estimate the business cycle involve constructing long-horizon conditional forecasts based on time series models. It has long been understood in the forecasting literature that combined forecasts can outperform individual forecasts (see, for example, Bates & Granger, 1969). Thus, combining model-based business cycle estimates could produce a measure with a lower mean-squared error than any of the individual estimates. Of course, the principle of combining forecasts does not answer the question of exactly how to combine forecasts. In this paper, we choose to use Bayesian model probabilities to construct weights on different forecasts.

Computation of equation (15) requires the posterior model probability, $\Pr(M_i|y)$, for each model. From Bayes' rule, this probability is proportional to the model's marginal likelihood multiplied by the prior model probability,

$$\Pr(M_i|y) \propto f(y|M_i)\Pr(M_i). \tag{16}$$

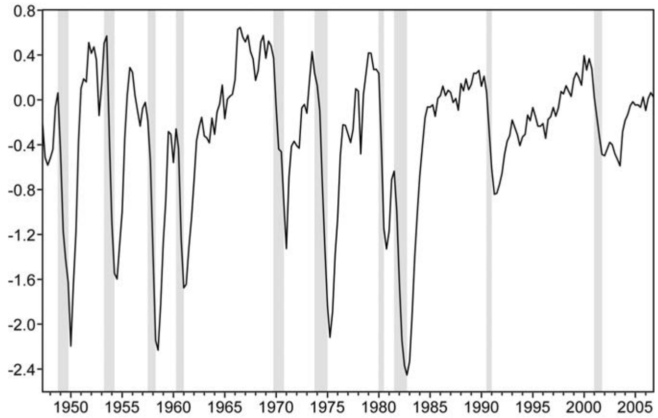
Direct calculation of the marginal likelihood, $f(y|M_i)$, requires averaging the likelihood function over all model parameters, where the averaging is done with respect to the prior distribution for the model parameters. This requires eliciting proper prior distributions for all parameters of each model under consideration, which can be extremely challenging for a large set of competing models. Complicating matters, marginal likelihood calculations are known to be sensitive to parameter prior specification, and models with diffuse parameter priors and many parameters are almost always dominated by models with diffuse parameter priors and fewer parameters. Thus, using nearly uninformative priors as a shortcut to avoid the difficult task of prior elicitation is generally not an option.

For the analysis here, we sidestep the need to elicit parameter priors by using an asymptotic approximation to the marginal likelihood of a model provided by the SIC statistic.[10] Under fairly general conditions, the SIC statistic is a consistent estimate of the log of the marginal likelihood. The advantage of the SIC statistic is that it relies on only maximum likelihood estimates and does not require elicitation of proper parameter priors. For this reason, as well as the relative ease of calculation, the SIC-based approximation is a popular choice in applied work.[11] Then, using the SIC statistic as an approximation to the log marginal likelihood, we have the following calculation for the posterior model probability:

$$\Pr(M_i|Y) = \frac{e^{SIC_i}\Pr(M_i)}{\sum_{i=1}^{N} e^{SIC_i}\Pr(M_i)}. \tag{17}$$

In addition to the SIC statistic, the posterior model probability depends on a prior model probability, $\Pr(M_i)$, for each model. In our analysis, we assign equal weight to the two classes of linear and nonlinear models. Then, within each class of models, we assign equal weight to each specification considered. There are 33 models under consideration, 15 of them linear and 18 of them nonlinear. Thus, each linear model receives prior weight proportional to 1/15, while each nonlinear model receives prior weight proportional to 1/18.

FIGURE 3.—MODEL-AVERAGED MEASURE OF THE U.S. BUSINESS CYCLE



NBER recessions are shaded.

The final columns of tables 2 and 3 report the posterior model probabilities constructed using the SIC-based approximation. Again, as discussed in section IV, essentially all posterior probability is accounted for by three groups of models that yield similar business cycle measures: bounceback models (40%), low-order AR models (44%), and linear UC models (14%). Figure 3 displays the model-averaged measure of the business cycle. Perhaps the most striking feature of this measure is its asymmetric shape, which it inherits from the bounceback models. In particular, the variation in the cycle is substantially larger during recessions than it is in expansions.

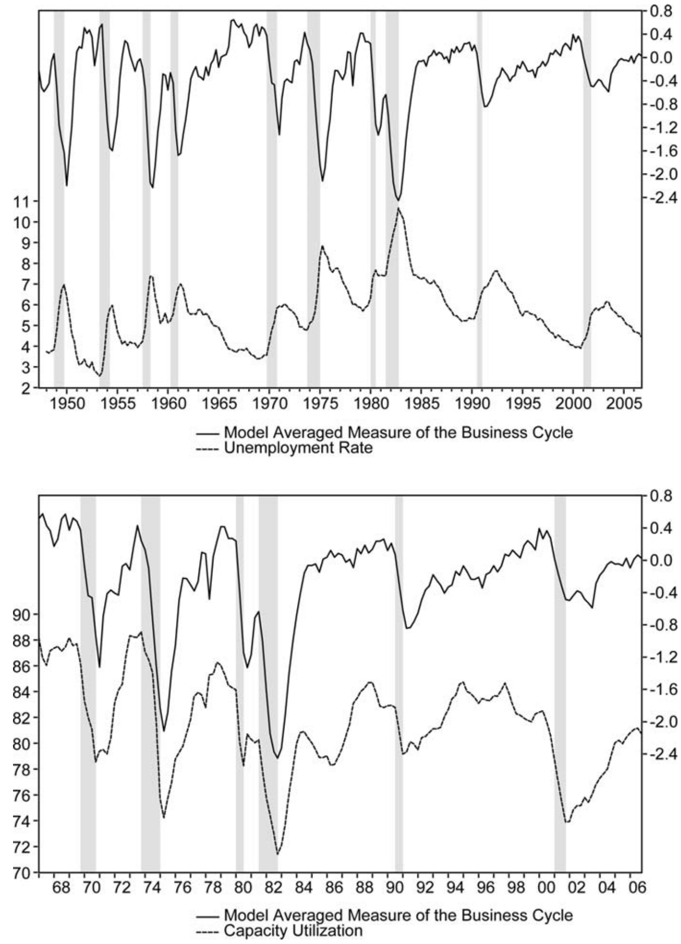
It is worth noting that this asymmetry is not a forgone conclusion given the weights on the bounceback models. Had the cycles implied by the preferred linear models all displayed substantial variation during expansions, then the model-averaged measure would have a more symmetric shape across business cycle phases. However, the low-order AR models, which receive the highest weight of the linear models, also display only small amounts of variability in expansions. Thus, when combined with the bounceback models, there is a total of 84% of the overall weight in the model-averaged measure given to measures of the cycle that display very little variation in expansion phases. Put differently, while there is substantial model-based uncertainty about the overall shape and magnitude of the cycle, there is relatively little model-based uncertainty about what the business cycle looks like during expansions.

The model-averaged business cycle measure is based on only the univariate dynamics of real GDP. Therefore, it is interesting to investigate its comovement with measures of economic slack based on other variables. Two variables that

[10] We have also constructed posterior model probabilities directly. To elicit parameter priors, we used a training sample of real GDP data to convert improper priors to proper priors and then constructed marginal likelihoods using the remainder of the sample. The model-averaged cycle resulting from this analysis was very close to that obtained using the SIC-based approximation. These results are available from us on request.

[11] See, for example, Brock, Durlauf, and West (2003) and Doppelhofer, Miller, and Sala-i-Martin (2004). For additional discussion of the SIC-based approach to model averaging, see Raftery (1995).

FIGURE 4.—COMPARISON OF MODEL-AVERAGED BUSINESS CYCLE TO OTHER MEASURES OF ECONOMIC SLACK



NBER recessions are shaded.

are often thought to correspond closely to the business cycle are the level of the unemployment rate and the level of capacity utilization. Figure 4 plots the model-averaged business cycle measure against these two variables.[12] There is a striking relationship between the model-averaged measure and both the unemployment rate and capacity utilization, with the short-run movements in these three series tracking each other quite closely. Of course, the advantage of the model-averaged measure as an indicator of the business cycle over these other readily available variables is that it is designed to capture transitory fluctuations in overall real economic activity while abstracting from all long-run variation. By contrast, while the unemployment rate and capacity utilization variables are thought to move with the business cycle, they are not as broad measures of economic activity as real GDP, and their historical paths suggest at least some permanent movements over time.

[12] The raw data for both variables are taken from the FRED database and cover the sample period 1948:Q1-2006:Q4 for the unemployment rate, and 1967:Q1-2006:Q4 for capacity utilization.

VII. Conclusion

We have provided estimates of the U.S. business cycle, where we define the business cycle as transitory deviations in economic activity away from trend. The estimates turn out to be highly dependent on the particular time series model used to capture postwar U.S. real GDP dynamics. As a result, we have attempted to discriminate between different implied business cycle measures by model comparison and formal hypothesis testing. The empirical results support a nonlinear regime-switching model that captures high-growth recoveries following deep recessions and produces a highly asymmetric business cycle with relatively small amplitude during expansions but large and negative movements during recessions. However, the model comparison also reveals several close competitors to the nonlinear model that produce business cycle measures of widely differing shapes and magnitudes. To address this model-based uncertainty, we constructed a model-averaged measure of the business cycle using posterior model probabilities as weights. We found that this model-averaged measure also displays strong asymmetry across NBER expansion and recession phases. Furthermore, the model-averaged business cycle is closely related to other measures of economic slack such as the unemployment rate and capacity utilization.

The asymmetry of the business cycle has many important implications. Most directly, it suggests a link between the output-gap definition of the business cycle and the NBER definition of the business cycle as alternating phases of expansion and recession. In particular, the asymmetry implies that NBER-dated recessions are periods of significant transitory variation in output, while output in expansions is dominated by movements in trend. This supports the idea that the business cycle is a meaningful macroeconomic phenomenon and has potential relevance as a stylized fact to guide theoretical models of the business cycle.

REFERENCES

Bai, Jushan, Robin L. Lumsdaine, and James H. Stock, "Testing for and Dating Common Breaks in Multivariate Time Series," *Review of Economic Studies* 65:3 (1998), 395–432.

Bates, John M., and Clive W. J. Granger, "The Combination of Forecasts," *Operations Research Quarterly* 20 (1969), 451–468.

Beveridge, Stephen, and Charles R. Nelson, "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* 7:2 (1981), 151–174.

Brock, William A., Steven N. Durlauf, and Ken D. West, "Policy Evaluation in Uncertain Economic Environments," *Brookings Papers on Economic Activity* 1 (2003), 235–301.

Carrasco, Marine, Liang Hu, and Werner Ploberger, "Optimal Test for Markov Switching," working paper (2007).

Clarida, Richard H., and Mark P. Taylor, "Nonlinear Permanent-Temporary Decompositions in Macroeconomics and Finance," *Economic Journal* 113:486 (2003), C125–C139.

Cogley, Timothy, and James M. Nason, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control* 19:1–2 (1995), 253–278.

Davidson, Russell, and James G. MacKinnon, *Econometric Theory and Methods* (New York: Oxford University Press, 2004).

Di Sanzo, Silvestro, "Testing for Linearity in Markov Switching Models: A Bootstrap Approach," *Statistical Methods and Applications* 18:2 (2009), 153–168.
Doppelhofer, Gernot, Ronald I. Miller, and Xavier Sala-i-Martin, "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review* 94:4 (2004), 813–835.
Garcia, René, "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review* 39:3 (1998), 763–788.
Hamilton, James D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57:2 (1989), 357–384.
——— "What's Real about the Business Cycle?" *Federal Reserve Bank of St. Louis Review* 87:4 (2005), 435–452.
Hansen, Bruce E., "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics* 7 (1992), S61–S82.
Harding, Don, and Adrian R. Pagan, "Dissecting the Cycle: A Methodological Investigation," *Journal of Monetary Economics* 49:3 (2002), 365–381.
——— "A Suggested Framework for Classifying the Modes of Cycle Research," *Journal of Applied Econometrics* 20:2 (2005), 151–159.
Harvey, Andrew C., and Albert Jaeger, "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics* 8:3 (1993), 231–247.
Kim, Chang-Jin, "Dynamic Linear Models with Markov Switching," *Journal of Econometrics* 60:1–2 (1994), 1–22.
Kim, Chang-Jin, James Morley, and Jeremy Piger, "Nonlinearity and the Permanent Effects of Recessions," *Journal of Applied Econometrics* 20:2 (2005), 291–309.
Kim, Chang-Jin, and Charles R. Nelson, "Friedman's Plucking Model of Business Fluctuations: Tests and Estimates of Permanent and Transitory Components," *Journal of Money, Credit and Banking* 31:3 (1999a), 317–334.
——— "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," this REVIEW 81:4 (1999b), 608–616.
Lam, Pok-sang, "The Hamilton Model with a General Autoregressive Component: Estimation and Comparison with Other Models of Economic Time Series," *Journal of Monetary Economics* 26:3 (1990), 409–432.
McConnell, Margaret M., and Gabriel Perez-Quiros, "Output Fluctuations in the United States: What Has Changed since the Early 1980s?" *American Economic Review* 90:5 (2000), 1464–1476.
Min, Chung-ki, and Arnold Zellner, "Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates," *Journal of Econometrics* 56:1–2 (1993), 89–118.
Mitchell, Wesley A., *Business Cycles: The Problem and Its Setting* (New York: National Bureau of Economic Research, 1927).
Morley, James, "A State-Space Approach to Calculating the Beveridge-Nelson Decomposition," *Economics Letters* 75:1 (2002), 123–127.
Morley, James, Charles R. Nelson, and Eric Zivot, "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP So Different?" this REVIEW 85:2 (2003), 235–243.
Morley, James, and Jeremy Piger, "Trend/Cycle Decomposition of Regime-Switching Processes," *Journal of Econometrics* 146:2 (2008), 220–226.
Morley, James, Jeremy Piger, and Pao-Lin Tien, "Reproducing Business Cycle Features: Do Nonlinear Dynamics Really Matter?" working paper (2010).
Murray, Christian J., "Cyclical Properties of Baxter-King Filtered Time Series," this REVIEW 85:2 (2003), 472–476.
Perron, Pierre, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica* 57:6 (1989), 1361–1401.
Perron, Pierre, and Tatsuma Wada, "Let's Take a Break: Trends and Cycles in U.S. Real GDP," *Journal of Monetary Economics* 56:6 (2009), 749–765.
Raftery, Adrian, "Bayesian Model Selection in Social Research," *Sociological Methodology* 25 (1995), 111–163.
Sichel, Daniel E., "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry* 31:2 (1993), 224–236.
——— "Inventories and the Three Phases of the Business Cycle," *Journal of Business and Economic Statistics* 12:3 (1994), 269–277.
Sinclair, Tara M., "Asymmetry in the Business Cycle: Friedman's Plucking Model with Correlated Innovations," *Studies in Nonlinear Dynamics and Econometrics* 14:1 (2009).
Wen, Yi, "The Business Cycle Effects of Christmas," *Journal of Monetary Economics* 49:6 (2002), 1289–1314.

APPENDIX A

Calculation of Trend Measures

This appendix provides formulas for the BN measure of trend implied by an autoregressive model, as well as for the RDSS measure of trend implied by the Hamilton and bounceback regime switching models.

Given a linear AR model of Δy_t, such as was given in equation (7), the BN trend for y_t can be easily calculated analytically using the state-space method in Morley (2002) as

$$\hat{\tau}_t^{BN} = y_t + HF(I-F)^{-1}(\Delta\tilde{y}_t - \tilde{\mu}), \tag{A1}$$

where

$$H = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}, F = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\Delta\tilde{y}_t = \left(\Delta y_t, \ldots, \Delta y_{t-p+1}\right)', \text{ and } \tilde{\mu} = (\mu, \ldots, \mu)'.$$

Then, for a nonlinear AR model, such as was given in equations (8) and (9), the RDSS trend can be calculated as

$$\hat{\tau}_t^{RDSS}\left(\tilde{S}_t\right) = y_t + HF(I-F)^{-1}(\Delta\tilde{y}_t - \tilde{\mu}_t) + \sum_{j=1}^{\infty}\left(E\left[\mu_{t+j}\middle|\{S_{t+k}=i^*\}_{k=1}^{j}, \tilde{S}_t, \Omega_t\right] - \bar{\mu}_{i^*}\right), \tag{A2}$$

where H, F, and $\Delta\tilde{y}_t$ are the same as in equation (A1), $\tilde{\mu}_t = \left(\mu_{S_t}, \ldots, \mu_{S_{t-m+1}}\right)'$, and $\bar{\mu}_{i^*} = \mu(i^*, \ldots, i^*)$. The first two terms on the right-hand side of equation (A2) are analogous to the calculation of the BN estimate of trend in equation (A1), with the second term corresponding to forecastable momentum due to linear dynamics. The third term arises due to nonlinear dynamics and accounts for forecastable momentum implied by any difference between the future time-varying mean and the regime-dependent average growth rate in regime i^*. When $i^* = 0$ for the models considered in this paper, the summation in equation (A2) can be truncated at $j = m$ and calculated analytically. In general, the infinite sum can be calculated via simulation.[13] Once $\hat{\tau}_t^{RDSS}\left(\tilde{S}_t\right)$ is calculated, $\tilde{S}_t$ can then be

[13] The RDSS approach is based on the assumption, discussed in Morley and Piger (2008), that permanent and transitory innovations depend on only current or lagged regimes. This assumption, which is explicitly made for the nonlinear UC models considered in this paper, implicitly holds for the BBU and BBV models, meaning that the specification of a "normal" regime affects only the level of the cycle, not its magnitude or shape. However, for the BBD model, the transitory effects of past shocks can depend on future regimes because the model has implicit regime-switching autoregressive coefficients. In this case, the level and magnitude of the cycle can be affected by the particular assumed sequence of future regimes, although the general shape will be robust. In additional analysis not reported here, we have also considered an extended simulation-based version of the RDSS approach that allows more complicated patterns for future regimes. Given the same "normal" regime at long horizons, we found nearly identical measures of the cycle, including in terms of magnitude to what is produced by the considerations of equation (A2) for the BBD model. Thus, for simplicity of presentation, we consider the basic RDSS approach for all of the bounceback models in this paper.

integrated out as in equation (5) using the probability weights, $p^M(\tilde{S}_t|\Omega_t)$, to arrive at $\hat{\tau}_t^{RDSS}$. These probability weights can be obtained from the recursive filter given in Hamilton (1989).

APPENDIX B

Implementation of the Nonlinearity Test

This appendix provides details on how to implement the CHP test developed by Carrasco et al. (2007) in the setting where the regime-switching alternative allows parameters to depend on both current and lagged values of the state variable. The CHP test considers the null hypothesis of constant parameters against an alternative hypothesis of switching parameters. Let θ_t denote the potentially time-varying parameters, a subset of all model parameters Θ. The null hypothesis is that H_0: $\theta_t = \theta_0$. The alternative hypothesis is that H_1: $\theta_t = \theta_0 + \theta_t^*$, where $\theta_t^* = H\tilde{\xi}_t$, which is the product of a matrix of possible changes in parameter values given switching regimes and $\tilde{\xi}_t$ is a vector of current and lagging 0 mean Markov state variables that determine the prevailing regime for the parameters. The state vector evolves according to $\tilde{\xi}_t = F\tilde{\xi}_{t-1} + w_t$, where the vector w_t follows a martingale difference sequence $\left(E\lfloor\tilde{\xi}_{t-1}'w_t\rfloor = 0\right)$, with $E\lfloor w_t w_t'\rfloor = Q$. Lagged state variables are incorporated into $\tilde{\xi}_t$ using identities, implying 0 elements in w_t.

The general form of the CHP test statistic is given as

$$TS_T(\beta) = \Gamma_T - \frac{1}{2T}\hat{\varepsilon}'\hat{\varepsilon}, \tag{B1}$$

where

$$\Gamma_T = \frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\gamma_t(\beta), \tag{B2}$$

$$\gamma_t(\beta) = tr\left(\left(l_{t,\theta}^{(2)} + l_{t,\theta}^{(1)}l_{t,\theta}^{(1)\prime}\right)E\left[\theta_t^*\theta_t^{*\prime}\right]\right) + 2\sum_{s<t} tr\left(l_{t,\theta}^{(1)}l_{s,\theta}^{(1)\prime}E\left[\theta_t^*\theta_s^{*\prime}\right]\right), \tag{B3}$$

$$l_{t,\theta}^{(1)} = \frac{\partial \ln f(y_t|\tilde{y}_{t-1},\Theta)}{\partial\theta}, l_{t,\theta}^{(2)} = \frac{\partial^2 \ln f(y_t|\tilde{y}_{t-1},\Theta)}{\partial\theta\partial\theta'}, \tag{B4}$$

and $\hat{\varepsilon}$ is the vector of residuals from an OLS regression of $1/2\gamma_t(\beta)$ on $l_{t,\Theta}^{(1)}$ (that is, the scores with respect to all of the parameters under the null, not just those that are hypothesized to switch under the alternative), with β denoting a vector of all of the nuisance parameters in H and F that are not identified under the null. Because of the presence of nuisance parameters, the test is based on the supremum test statistic for a set of considered values of the nuisance parameters (a "sup" test statistic $\sup TS = \sup_{\beta\in\bar{B}} TS_T(\beta)$, where $\bar{B}$ is a compact subset of all possible values of the nuisance parameters B). Note that for the sup statistic, the scale of the nuisance parameters in H is not identified because it cancels out in the first-order condition with respect to that scale parameter. Thus, the value of an arbitrary nuisance parameter that is assumed not to take on the value of 0 under the alternative can be normalized to 1 and the test statistic can be constructed as

$$\sup TS = \sup_{\beta'\in\bar{B}'}\frac{1}{2}\left(\max\left(0,\frac{\Gamma_T}{\sqrt{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}}\right)\right)^2, \tag{B5}$$

where β is a vector of the remaining nuisance parameters and $\hat{\varepsilon}^* = \hat{\varepsilon}/\sqrt{T}$. Also, note that the expectations terms in equation (B3) can be solved as $E\lfloor\theta_t^*\theta_t^{*\prime}\rfloor = H\mathrm{Var}\left(\tilde{\xi}_t\right)H'$, where

$$vec(\mathrm{Var}\left(\tilde{\xi}_t\right)) = (I - F\otimes F)^{-1}vec(Q),$$

$$\text{and } E\lfloor\theta_t^*\theta_s^{*\prime}\rfloor = HF^{t-s}\mathrm{Var}\left(\tilde{\xi}_t\right)H'.$$

The asymptotic distribution of the CHP test depends on nuisance parameters. As a result, Carrasco et al. (2007) rely on parametric bootstrap experiments to calculate the critical values. These experiments involve simulating B bootstrap samples based on the estimated null model and calculating the test statistic for each of these simulated samples. Then the percentage of simulated test statistics larger than the sample statistic determines the bootstrap p-value for the test, while the bootstrap critical value for a test with nominal size α can be found by sorting the bootstrap test statistics from smallest to largest and finding the $(1-\alpha)B$ test statistic or the next largest if $(1-\alpha)B$ is not an integer.

A simple example helps illustrate the CHP test. Consider the null of an AR(0) model $\Delta\gamma_t = \mu + e_t$, $e_t \sim N(0,\sigma_e^2)$ against the alternative of a two-state Markov-switching mean $\Delta\gamma_t = \mu_t + e_t$, where $\mu_t = \gamma_0 + \gamma_1 S_t$ and $S_t = \{0,1\}$ is a two-state Markov-switching state variable with fixed continuation probabilities $\Pr[S_t = 0|S_{t-1} = 0] = p_{00}$ and $\Pr[S_t = 1|S_{t-1} = 1] = p_{11}$. Then, letting $\bar{\pi} \equiv E[S_t] = (1-p_{00})/(2-p_{11}-p_{00})$, $\mu_t = \mu + \gamma_1\xi_t$, where $\mu = \gamma_0 + \gamma_1\bar{\pi}$ and $\xi_t \equiv S_t - \bar{\pi}$. Thus, in terms of the general CHP test, $H = \gamma_1$, $\tilde{\xi}_t = \xi_t$, and $F = \rho$, where $\rho = p_{00} + p_{11} - 1$. Note that it is necessary to normalize the variance of the unobserved state variable in order to identify the magnitude of γ_1. We do this by setting $Q = 1$. Then, in constructing the test statistic, we set $\gamma_1 = 1$ and find the largest test statistic for $\rho \in (0.02, 0.98)$. In practice, we consider only positive values for ρ because alternatives with persistent regimes are what we are interested in for U.S. real GDP. Specifically, we are considering regime-switching models in order to capture persistent business cycle phases rather than outliers. The restriction on ρ can lower the value of the test statistic, and in some cases it does. However, the same restriction is imposed when calculating the bootstrap distribution of the test statistic, so it can also lower the critical value of the test. Also, it is worth mentioning that we want to avoid the case where $\rho = 0$ because given the assumption of a linear Gaussian model under the null hypothesis, the second derivatives in equation (B3) are equal to the negative of the outer product of the scores $\left(l_{t,\theta}^{(2)} = -l_{t,\theta}^{(1)}l_{t,\theta}^{(1)\prime}\right)$, meaning that $\gamma_t(\beta) = 0$ and therefore making it impossible to run the OLS regression to find $\hat{\varepsilon}$. As discussed in Carrasco et al. (2007), the test has no power in this case. On the other hand, the test has nontrivial power in other cases, including, in this setting, when $\rho \neq 0$.