Wesleyan Economic Working Papers

http://repec.wesleyan.edu/
N^{o}: 2009-003

Reproducing Business Cycle Features: How Important Is Nonlinearity Versus Multivariate Information?

James Morley, Jeremy Piger, and Pao-Lin Tien

May 2009

WESLEYAN UNIVERSITY

Department of Economics
Public Affairs Center
238 Church Street
Middletown, CT 06459-007

Tel: (860) 685-2340
Fax: (860) 685-2301
http://www.wesleyan.edu/econ

Reproducing Business Cycle Features: How Important Is Nonlinearity Versus Multivariate Information?

James Morley, Jeremy Piger, and Pao-Lin Tien*

This draft: May 11th, 2009

Abstract

In this paper, we consider the ability of time-series models to generate simulated data that display the same business cycle features found in U.S. real GDP. Our analysis of a range of popular time-series models allows us to investigate the extent to which multivariate information can account for the apparent univariate evidence of nonlinear dynamics in GDP. We find that certain nonlinear specifications yield an improvement over linear models in reproducing business cycle features, even when multivariate information inherent in the unemployment rate, inflation, interest rates, and the components of GDP is taken into account.

JEL Classification: E30, C52

* James Morley, Washington University in St. Louis; Jeremy Piger, University of Oregon; Pao-Lin Tien, Wesleyan University.

1. Introduction

Model evaluation has always been at the forefront of macroeconomic research. As modeling techniques have advanced over time, a wide variety of time-series models have sprung up to satisfy different needs, from simple univariate and multivariate linear models to more complicated univariate and multivariate nonlinear models. It is therefore important to establish an efficient and reasonable approach to model comparison and evaluation that is suitable for very different types of time-series models. In this paper, we evaluate a variety of univariate linear, univariate nonlinear, and multivariate linear models of U.S. real gross domestic product (GDP) in terms of their abilities to produce simulated data that exhibit the business cycle features in the actual GDP data. Our primary goal is to investigate the extent to which multivariate information inherent in macroeconomic variables such as the unemployment rate, inflation, interest rates, and the components of GDP can account for the apparent univariate evidence of nonlinear dynamics in U.S. GDP previously demonstrated in the literature.

The conventional methods for conducting model evaluation – hypothesis testing, out-of-sample forecast comparisons, and Bayes factors – face several drawbacks. When the models under consideration are non-nested, hypothesis testing is often intractable. Out-of-sample forecast comparisons tend to be sensitive to the particular out of sample period used. Bayes factors can be very sensitive to the specification of priors. Furthermore, Bayes factors only provide a sense of the relative performance of different models and not an absolute measure of the ability of a model to explain the dynamics in the data.

The business cycle features approach considered in this paper offers a useful alternative to the conventional methods for model evaluation. It can be viewed as being related to a broader approach to model comparison and evaluation known as "encompassing tests." Encompassing tests evaluate the ability of models to produce simulated data that have the same behavior as sample data. In our particular case, we concentrate on features of the data that are related to business cycles. Ever since Burns and Mitchell's (1947) extensive study of the cyclical behavior of economic activity, economists have sought to analyze economic fluctuations in terms of business cycle phases. Thus, our focus on business cycle features provides a very natural way to assess the benefit of introducing nonlinearity into time-series models, especially because many of the nonlinearities explored for GDP have been motivated as being related to the business cycle.

One can also view the encompassing method of model evaluation as complementary to the more traditional methods. For example, if several non-nested models – such as an ARIMA model and a Markov-switching model – manage to pass the battery of conventional diagnostic tests and no particular model dominates all others in terms of different out-of-sample periods, then these models' abilities to produce simulated data that can match the business cycle features of GDP could help researchers choose which model is most useful in the context of analyzing the business cycle.

We employ the business cycle features approach to compare the preferred univariate linear and nonlinear models in Morley and Piger (2006) with three popular multivariate linear models: the two-variable vector autoregression (VAR) model of Blanchard and Quah (1989); the four-variable VAR model in Ahmed, Levin, and Wilson (2004); and the three-variable vector error correction model (VECM) in King, Plosser, Stock, and Watson (1991). What we find is that multivariate information does not appear to improve the performance of linear models over nonlinear models. These results are robust even when a structural break in the variance of U.S. real GDP is taken into account. Also, we find no clear advantage to using non-parametric versus parametric residual specifications for data simulation. These results strengthen the argument that certain parametric nonlinear specifications capture business cycle features of the data that no linear models or nonparametric residuals can explain.

The remainder of this paper proceeds as follows: Section 2 provides a brief overview of the literature on business cycle features; Section 3 details the business cycle algorithm used to establish business cycle turning points in U.S. real GDP; Section 4 defines the business cycle features that we consider and documents these features for U.S. real GDP; Section 5 specifies the time-series models under consideration and then evaluates the abilities of the competing univariate and multivariate models to reproduce business cycle features exhibited by GDP; Section 6 concludes.

2. Literature Review

A number of recent papers in the literature have employed the business cycle features approach to assess the performance of different time-series models, including Hess and Iwata (1997), Harding and Pagan (2002), Galvão (2002), Clements and Krolzig (2004), and Morley and Piger (2006) for U.S. data, and Demers and Macdonald (2007) for Canadian data.

In the plethora of univariate and multivariate linear and nonlinear models that Hess and Iwata (1997), Harding and Pagan (2002), and Clements and Krolzig (2004) considered, the simple linear ARIMA(1,1,0) or ARIMA(2,1,0) models always manage to reproduce business cycle features of actual real GDP just as well as, if not better than, their more complicated counterparts. Following the principle of parsimony, all three papers draw the conclusion that researchers should pick the simpler models over more complicated models, ceteris paribus. However, Galvão (2002), Morley and Piger (2006), and Demers and Macdonald (2007) find that, while none of the models being considered dominates over all features, there are some important features that certain nonlinear models are better at capturing than linear models. Hence there is added benefit and relevance for taking into account nonlinearity in time-series models.

Of the above mentioned papers in the literature, only Clements and Krolzig (2004) have systematically compared univariate models against multivariate models, and they find that multivariate models do not do very well in terms of matching business cycle features of U.S. real GDP. However, in this paper, we consider a set of business cycle features that differ from those in Clements and Krolzig (2004). We include not only the typical growth rates and durations of different business cycle phases, but also consider the correlation between the cumulative growth rate observed during recessions and that observed in the subsequent recovery phase, a characteristic of U.S. GDP data that was central to Milton Friedman's (1964, 1993) analysis of the U.S. business cycle. In addition, Clements and Krolzig (2004) used a business cycle dating algorithm that does not impose a minimum length requirement for business cycle phases. The business cycle dating algorithm (MBBQ) we implement in this paper does impose some minimum length requirements, and this algorithm has been shown to produce peak and trough dates that match the National Bureau of Economic Research's (NBER) business cycle dates better than the commonly adopted dating algorithm (BBQ). Finally, even though we do not consider as many multivariate models as Clements and Krolzig (2004), the models we do look at are some of the most popular and widely used multivariate linear models, and we also allow for more flexibility in terms of residual distribution for data simulation purposes.

3. Business Cycle Dating Algorithm

Before discussing business cycle dating algorithms, we must first define what we mean by the business cycle. Under the business cycle features comparisons framework, "cycle" refers to the *classical* business cycle (or *reference cycle*) as described by Burns and Mitchell (1947) rather than the *cyclical component* of a series obtained after detrending the data series, although the two concepts may be closely related (see Morley and Piger 2009). According to Burns and Mitchell (1947), the business cycle can be defined as a series of distinct phases in economic activity, with the phases corresponding to recession and expansion. The turning points of the phases are indicated as peaks and troughs. The de facto business cycle peak and trough dates in the U.S. are determined by the NBER, a private, nonprofit, nonpartisan research organization founded in 1920. Within the NBER, the Business Cycle Dating Committee plays the key role in establishing business cycle dates. The committee reviews a variety of economic statistics and indicators of U.S. business conditions before deciding on the exact turning points in the economy.

The NBER business cycle dates are widely used in economic research requiring business cycle peak and trough dates, and it seems natural to use them as the benchmark for calculating business cycle features. However, the NBER chronology is only relevant for the actual U.S. sample data, and not for the simulated data from the time-series models we are considering. Therefore, to establish turning points in the sample data and simulated data in a consistent fashion, we need to use a formal procedure capable of mimicking the NBER decision-making process.

The standard approach to establishing business cycle turning points in the literature is to use the Bry-Boschan Quarterly (BBQ) algorithm developed by Harding and Pagan (2002). This is a quarterly version of the BB algorithm for monthly data proposed by Bry and Boschan (1971). The specifics of the algorithm can be summarized as follows:

Step 1: Using the log level of U.S. quarterly real GDP (y_t), establish candidate dates of peaks and troughs as local maxima and minima in the data such that a peak occurs at time t if:

$y_{t-2} - y_t < 0; \quad y_{t-1} - y_t < 0; \quad y_{t+1} - y_t < 0; \quad y_{t+2} - y_t < 0,$

and a trough occurs at time t if:

$y_{t-2} - y_t > 0; \quad y_{t-1} - y_t > 0; \quad y_{t+1} - y_t > 0; \quad y_{t+2} - y_t > 0.$

Step 2: Censor the turning points to ensure that peaks and troughs alternate. In the case of two consecutive peaks (troughs), eliminate the peak (trough) with the lower (higher) value of y_t.

Step 3: Censor the turning points to ensure that each business cycle phase (peak-to-trough and trough-to-peak) lasts a minimum of two quarters, while each complete business cycle (peak-to-peak and trough-to-trough) lasts a minimum of five quarters.

The peak and trough dates established by the NBER for the sample period 1948Q4 to 2007Q4,[1] along with the dates established by the BBQ algorithm applied to quarterly U.S. real GDP are reported in Table 1. The BBQ algorithm does a reasonable job of matching the NBER peak and trough dates. It identifies eight of the nine peaks and nine of the ten troughs reported by

[1] Even though U.S. real GDP data are available as early as 1947Q1, we choose to start our sample at 1948Q4. As a result, we have to ignore the first NBER peak date (1948Q4) in our analysis, as the earliest start date at which the dating algorithms can identify a turning point is 1949Q2. The main reason for starting the sample at 1948Q4 is that starting the sample at 1947Q1 creates problems for the dating algorithms considered here. It not only causes the BBQ algorithm to pick up an extra trough date in 1947Q3, but it also throws off the precision of the dating algorithms in terms of their ability to produce trough dates that match those reported by the NBER. We believe that this is due to the interaction of the 1947Q1 observation with the minimum phase length and censoring requirements in Steps 2 and 3 of the algorithms. We consider shortening the sample period by seven quarters to be a worthwhile sacrifice in order to make the dating algorithms more precise.

NBER. Just two of the peak dates differ from the corresponding NBER peak dates, each by a single quarter, while five of the trough dates differ from the corresponding NBER trough dates, with the differences ranging from one to three quarters. It is interesting that all the errors made by the BBQ algorithm shift the turning points forward in time relative to the NBER dates. This systematic error suggests that Step 1 of the BBQ algorithm can be modified to correct for it.

Morley and Piger (2006) modified the BBQ algorithm by optimizing on the threshold values that indicate turning points. We refer to this modified BBQ algorithm as MBBQ. Specifically, MBBQ restates Step 1 of the BBQ algorithm as follows:

Step 1: Using the log level of U.S. quarterly real GDP (y_t), establish candidate dates of peaks and troughs as local maxima and minima in the data such that a peak occurs at time t if:

$$y_{t-2} - y_t < \alpha_1; \quad y_{t-1} - y_t < \alpha_1; \quad y_{t+1} - y_t < \alpha_2; \quad y_{t+2} - y_t < \alpha_2,$$

and a trough occurs at time t if:

$$y_{t-2} - y_t > \alpha_3; \quad y_{t-1} - y_t > \alpha_3; \quad y_{t+1} - y_t > \alpha_4; \quad y_{t+2} - y_t > \alpha_4.$$

MBBQ differs from BBQ in that the threshold parameters that signal turning points are allowed to deviate from 0. The thresholds are also allowed to vary from peak to trough and on different sides of the turning points. To determine the values of the α_i's, i = 1, 2, 3, 4, a grid search is conducted for values between –0.005 and 0.005, i.e. $\alpha_i \in (-0.005, 0.005)$. For each possible combination of the α_i's in the grid, the root mean squared error (RMSE) is calculated as:

$$RMSE(\alpha_i) = \sqrt{\frac{\sum_{t=1}^{T} [MBBQ_t(\alpha_i) - NBER_t]^2}{T}},$$

where $NBER_t$ = 1 if quarter t is an NBER recession quarter and $NBER_t$ = 0 otherwise, while $MBBQ_t(\alpha_i)$ = 1 if quarter t is a recession quarter according to the MBBQ algorithm with threshold values α_i, and $MBBQ_t(\alpha_i)$ = 0 otherwise. The α_i's that minimize RMSE(α_i) are chosen to be the final threshold values for the algorithm. In the case of ties, α_i's that are closest to 0, as measured by $\sum_{i=1}^{4}|\alpha_i|$, are chosen.

The turning point dates established by the MBBQ algorithm are reported in Table 1 as well. Threshold values chosen for this sample period are: $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0.001$, $\alpha_4 = -0.002$. It is clear from Table 1 that the MBBQ algorithm offers substantial improvement over the BBQ algorithm, especially on the trough dates. It identifies the same number of peaks and troughs as the BBQ algorithm, though only two of the peak dates and two of the trough dates deviate from their corresponding NBER dates, each by a single quarter.

Note that both the BBQ and MBBQ algorithms miss the peak and trough dates identified by the NBER in 2001. This was not the case in Morley and Piger (2006). Upon closer inspection of the data, we found that due to a benchmark data revision in 2004, the U.S. real GDP output growth rate for 2001Q2 was changed from negative to positive.[2] As both dating algorithms require two quarters of decline or increase on both sides of turning points, this revision in GDP data implies that neither algorithm would be able to pick up any peaks or troughs in 2001. The data revision hence diminishes the ability of the dating algorithms to mimic actual NBER chronology. However, given that both BBQ and MBBQ still do fairly well in picking out turning point dates that match up with the NBER dates prior to 2001, we believe that this problem is not serious enough for us to abandon the use of these algorithms.[3]

---

[2] According to the St. Louis Fed Archival Federal Reserve Economic Data (ALFRED), U.S. real GDP (GDPC1) with a vintage date of June 25th 2004 still reports a negative growth rate for 2001Q2, but in the next vintage (July 30th 2004) the same growth rate is revised to a positive number.

[3] There is ample evidence that 2001 remains a recession phase despite the revision in GDP data. In a recent memo released on January 7th, 2008 by the NBER Business Cycle Dating Committee, there is no mention of possibly revising the 2001 peak and trough dates. Also, even though the 2001 recession is no longer obvious from the level of the GDP series alone, it is still apparent in other series such as employment (total nonfarm payroll). In addition, nonlinear Markov-switching type models like the Kim et. al. (2005) bounceback model that we consider here still identify 2001 as a recession episode with the updated GDP data. Another interesting anecdote is that if we feed real gross domestic income (real GDI) into the algorithms rather than real GDP, both BBQ and MBBQ pick up the 2001 peak and trough dates, although they miss the 1980 peak and trough instead. Hence, despite the recent attention paid to GDI by the Business Cycle Dating Committee in their most recent report on the determination of the December 2007 peak in economic activity, using GDI does not offer an absolute improvement to using GDP in terms of producing peak and trough dates that match the NBER dates.

4. Business Cycle Features in U.S. Real GDP Data

The business cycle phases are defined as follows: (1) Recession – the quarter following a peak date to the subsequent trough date, (2) Expansion – the quarter following a trough date to the subsequent peak date, (3) Recovery – the first four quarters of the expansion phase, and (4) Mature Expansion – the remaining quarters of an Expansion phase following the Recovery phase.

Given this definition of phases, we consider the following business cycle features for any given realization of data:

- Number of business cycle peaks
- Average and standard deviation of Recession and Expansion phase lengths
- Average and standard deviation of annualized quarterly growth rates in Recession, Expansion, Recovery, and Mature Expansion phases
- Correlation between the cumulative decline during a Recession and the cumulative growth in the subsequent Recovery phase.

Table 2 presents the values of these business cycle features for quarterly U.S. real GDP data from 1948Q4 to 2007Q4 using turning points established by the NBER, the BBQ algorithm, and the MBBQ algorithm. The results here corroborate what we observe in Table 1; specifically, MBBQ does a better job at matching the NBER sample feature values than BBQ because it is better able to replicate NBER turning points. In all but four cases (average quarterly growth rates of the Expansion phase, average length of the Expansion phase, and the variation in the average length of Recession and Expansion phases) MBBQ produce feature values that are closer to the NBER sample features. Hence, in the following sections, we will compare the simulated features using the MBBQ algorithm against the NBER sample features.[4]

[4] Alternatively, we could have compared the simulated features against the sample features produced by the dating algorithm. However, due to complications with missing the 2001 peak and trough dates, there are some large differences between the sample features generated using the NBER turning points and those produced by MBBQ. Because the time-series models are designed to replicate the behavior of actual GDP with NBER recessions and expansions, we chose to compare simulated features with the NBER sample features instead.

Before proceeding to the discussion of features for simulated data in the next section, there are a few things worth mentioning regarding the NBER sample features reported in Table 2. First, as one would expect, average quarterly growth rates differ quite a bit between the Recession and Expansion phases. Recessions are associated with negative growth rates, averaging around –1.9% per quarter, while Expansions are associated with positive growth rates close to 4.6% per quarter.[5] When the Expansion phase is divided up into Recovery and Mature Expansion phases, it is striking to see that the average growth rate associated with the Recovery phase is almost twice as large as those reported for the Mature Expansion phase. Second, there is a large difference between the average length of the Recession and Expansion phases. Expansions appear to last nearly six times as long as Recessions. Third, the variability associated with the Recovery phase is much higher than for other phases in terms of the average quarterly growth rates. This high variability also applies to the average length of Expansion phase. Finally, there is strong negative correlation between the cumulative growth in a Recession phase and the cumulative growth in the subsequent Recovery phase. This corroborates the observation made in Friedman (1964, 1993).

5. Business Cycle Features in Simulated Data from Time-Series Models

5.1. Univariate Model Description

Two different univariate models are considered in this paper. First is the linear AR(2) model that has been found to do quite well in terms of matching business cycle features in the literature, and is the preferred model in Clements and Krolzig (2004). Second is the Kim, Morley, and Piger (2005) bounceback model, which is a nonlinear model with Markov-switching parameters. This version of the bounceback model is termed BBV indicating that this particular specification will be able to depict V-shaped recessions.[6] The key difference between the

[5] All growth rates are expressed in annualized terms.

[6] V-shaped recession refers to recessions exhibiting "sharpness," a term introduced by McQueen and Thorley (1993). A sharp series has the transition from contraction to expansion occurring more rapidly than the transition from expansion to contraction. This feature results in the level series being more rounded at peaks than at troughs.

bounceback model and the standard Hamilton (1989) two-state Markov-switching model is that it would be able to capture a high-growth recovery phase following the end of recessions. Furthermore, the strength of this high-growth recovery phase is related to the severity of the previous recession, as measured by its length. The BBV was the best performing time-series model in Morley and Piger (2006), beating even the three-state Markov-switching model of Boldin (1996), which was also designed to capture high-growth recovery business cycle phases.

The specification and estimates of the two time series models for quarterly U.S. real GDP are presented in the appendix. The reported estimates are what we used to calibrate the data generating process in our Monte Carlo simulations that will be used for business cycle feature comparisons later on.

5.2. Multivariate Model Description

As mentioned in the introduction, we consider three different multivariate models. The two-variable VAR model of Blanchard and Quah (1989) (B&Q), the four-variable VAR model in Ahmed, Levin, and Wilson (2004) (ALW), and the three-variable VECM in King, Plosser, Stock, and Watson (1991) (KPSW). The specifications and estimates used for the Monte Carlo simulations of the multivariate linear models are presented in the appendix.[7] These three models are of particular interest to us because they are widely cited multivariate models in the economics literature, and are specifically designed to explain aggregate economic fluctuations

In Blanchard and Quah (1989), the authors looked at the dynamic effects of aggregate demand and supply disturbances on gross national product (GNP) by using GNP growth and the unemployment rate in their VAR system. In Ahmed, Levin, and Wilson (2004), the authors

[7] Data used for estimation of the multivariate models vary from those used in the original papers. If the original model selected an output variable that was not real GDP (for example, Blanchard and Quah 1989 used real gross national product), we replace it with real GDP in our estimation. As for the other variables in the models, we try to stay as close to those in the original study as possible. The estimation sample periods for the multivariate models all start somewhat later than 1948Q4 for a variety of reasons (B&Q sample starts from 1950Q1, ALW starts from 1955Q3, and KPSW starts from 1949Q2), sometimes it is due to data availability, sometimes it is because of the number of lags the estimation requires, and sometimes it is both. We try to implement the longest possible sample to obtain the parameter estimates. Note, however, that we continue to simulate data from 1948Q4 to 2007Q4 even though the estimated parameters are based on different sample periods.

investigated the source of the reduction in the volatility of GDP growth since 1984, and in their VAR system they included GDP growth, inflation, commodity price inflation and the federal funds rate. A very similar VAR model to that used in Ahmed, Levin, and Wilson (2004) was also implemented in Stock and Watson (2002) and Boivin and Giannoni (2006). The VECM in King et. al. (1991) is a classic model for looking at the importance of productivity shocks on economic fluctuations. The authors claim that their analysis applies to a wide class of real business cycle models and is superior to the bivariate VAR in Blanchard and Quah (1989). They included private GNP (y), consumption (c), and investment (i) in their system with $(c-y)$ and $(i-y)$ as theory-based error-correction terms.[8]

In the next two subsections we use the estimated parameters reported in the appendix to simulate artificial real GDP series from 1948Q4 to 2007Q4, using the actual value of real GDP in 1948Q4 as an initial value. For each model, we perform 10,000 simulations, saving the business cycle features for each simulation. Following the convention in the literature, we neglect parameter uncertainty in our simulations. Thus, the only source of variation across simulations arises from the residuals, which, in most of the literature, are assumed to be normally distributed. However, we also consider whether this parametric specification for the residuals might be improved upon by using a semi-parametric bootstrap approach – that is, we shuffle the original residuals from the model estimation and then draw from this pool of residuals with replacement in order to construct the simulated series. This is a more general approach as no distributional assumptions about the residuals are being made (i.e. the residuals are non-parametric). If the true residuals are not normally distributed, the semi-parametric bootstrap approach should improve the performance of the models in terms of the simulated data's ability to reproduce business cycle features. This also helps us address any concerns that nonlinear models might be better than linear models at replicating business cycle features only because they can capture fat tails or skewness in the unconditional distribution of output growth rather than because of nonlinear dynamics inherent in the models.

[8] In the structural VAR literature, the type of identification method used is of vital importance. Blanchard and Quah (1989) and King et. al. (1991) implemented long-run restrictions while Ahmed, Levin, and Wilson (2004) used short-run restrictions. However, for the purpose of simulating data and calculating business cycle feature required here, identification of structural shocks is irrelevant. What matters are the variables included in each VAR or VECM model and the reduced-form dynamics generated by the models.

5.3. Business Cycle Features of Univariate Models

Table 3 presents the medians of the simulated distribution of each business cycle feature we consider for the univariate models. The median value for each feature is followed by (in parentheses) the proportion of simulations that fall below the corresponding sample feature reported in column 1 of the table for actual real GDP using the NBER turning point dates. These percentiles provide us with a sense of how likely the univariate models could have produced a sample value for a particular business cycle feature as large or small as that exhibited by the actual GDP data. Percentiles that are less than 0.10 or greater than 0.90 are bolded to show that it was highly unlikely that the particular univariate time-series model could have simulated data that replicates the behavior of actual GDP for that particular feature. The reported medians give us a sense of whether a percentile is driven by closeness of the distribution in matching the sample feature or by a large dispersion of the simulated distribution.

As an example, consider the number of peaks feature for the AR(2) model with parametric residual draws (second column of Table 3). The NBER reports 9 peaks, and the median of the simulated distribution for this feature is 9, indicating that the median of the 10,000 simulations from the AR(2) model produced 9 peaks. At the same time, 40% of the 10,000 simulations produced a number of peaks below 9, and 60% of the 10,000 simulations produced a number of peaks equal to or above 9. So we can interpret the results here as suggesting that the AR(2) model with parametric residual draws do a reasonably good job matching this particular feature exhibited by actual U.S. real GDP.

For the other business cycle features considered, the AR(2) model with parametric residual draws does a satisfactory job matching the numbers reported for actual U.S. real GDP. It is particularly good at replicating the features related to the number or length of phases. However, the large difference between the median value in the simulated data and the sample value for the average length and standard deviation of the Expansion phase shows that there is substantial dispersion in the simulated distribution. Also, the AR(2) model fails to reproduce the high Recovery growth rates exhibited by real GDP, and the standard deviation of quarterly growth rates for the phases are very far off from the sample data values. Finally, the AR(2) model does a very poor job at replicating the strong negative correlation between the cumulative growth rates of the Recession and Recovery phases exhibited by actual GDP. As column 3 demonstrates, similar results are obtained for the AR(2) model when non-parametric residuals are used in the simulations.

Turning our attention to the bounceback BBV model, we can see that it clearly fares better than the AR(2) model. Column 4 of Table 3 shows that the BBV model with parametric residual specification can match all features reasonably well except for the standard deviation of quarterly growth rates of Recessions. It is especially notable that the BBV model can capture the high quarterly growth rate during the Recovery phase as well as the strong negative correlation between the cumulative growth rate in the Recession phase and the cumulative growth rate in the subsequent Recovery phase. Non-parametric residuals in this case do not lead to an improvement in the performance of the BBV model at all, creating percentiles in excess of 0.9 for the average quarterly growth rates of Recession and Expansion phases. However, they do allow the BBV to generate a slightly stronger negative correlation between the cumulative growth during Recession and Recovery phases.[9]

The results reported here are consistent with the findings in Galvão (2002), Morley and Piger (2006), and Demers and Macdonald (2007) that certain nonlinear univariate models do a better job at capturing important asymmetries in the business cycles that are missed by linear univariate models.

5.4. Business Cycle Features of Multivariate Models

Table 4 reports the simulated business cycle features of multivariate models. A brief glance at the table reveals that the three different multivariate models produce more or less the same results. All three models do well in terms of matching the number of peaks and the average

[9] The weaker performance of the BBV model with bootstrapped residuals could be due to the problem of measuring residuals for such a model. In particular, residuals for Markov-switching models cannot be directly observed as they depend on the state (recession or expansion) and probability of switching or staying in that state. To get around this problem, we assume the state is observable by imposing the NBER peak and trough dates. Then, with the estimated model parameters, we calculate a set of residuals based on these states, allowing us to carry out the semi-parametric bootstrap procedure for the simulation exercise.

length and variation of Recession and Expansion phases. However, as with the linear AR(2) model earlier, they fail completely in terms of being able to generate a high enough average quarterly growth rate for the Recovery phase or a strong enough negative correlation between the cumulative growth rate of Recession and the cumulative growth rate of Recovery phases. The ALW four-variable VAR model even has trouble with the average quarterly growth rates in the Expansion phases. The multivariate models also cannot replicate the standard deviations associated with the quarterly growth rates of most of the business cycle phases.

Switching from parametric residual to non-parametric residuals improves the performance of all the multivariate models slightly. Mostly the improvement can be seen in being better able to match the standard deviation of quarterly growth rates of the business cycle phases. Consistent with the univariate findings, non-parametric residuals also help with generating a slightly stronger negative correlation between the cumulative growth rates of Recession and Recovery phases, though not strong enough to push the percentiles into an acceptable range. For the KPSW model, the non-parametric residuals actually worsen the performance of the model somewhat, by generating a median value of the average Expansion quarterly growth rate that is far too small relative to the actual real GDP sample value.

Given the results reported in Table 4, one can conclude that multivariate information does not improve the performance of linear models at replicating business cycle features of real GDP. In the best case scenario, the B&Q model with non-parametric residuals replicate features about as well as the simple AR(2) with non-parametric residuals. This result is quite consistent with that reported in Clements and Krolzig (2004), who find multivariate models do no better, and often worse, than the univariate linear ARIMA models.

So far, we have shown that the bounceback BBV model is the best performing model. However, it is important to note that not all nonlinear time-series models do better in terms of business cycle feature reproduction when compared to linear models. For example, Morley and Piger (2006) found that the two-regime Markov-switching model of Hamilton (1989) performs about the same as the linear models. A key reason why the nonlinear BBV model does a superior job in reproducing business cycle features is that there is a mechanism embedded in the model to capture high growth recoveries. This is what Galvão (2002) found as well when considering related models. Among the fifteen univariate nonlinear models she investigated, only two (a three-regime Markov-switching model and an unobserved components model with Markov-switching in the transitory component) were able to account for the asymmetries in the shape of the U.S. business cycle, and those two models are both characterized by mechanisms to capture high growth recoveries.

5.5. Business Cycle Features and the "Great Moderation"

There is much evidence for a marked decline in the volatility of U.S. real GDP growth since the mid 1980s, which is often labeled the "Great Moderation." The magnitude of the decline is striking. McConnell and Perez-Quiros (2000) show that the variance of output fluctuations since 1984 is only one fourth of the variance for the period ending in 1983. There is much debate about the reason for the decline in volatility; some argue it is good monetary policy or better business practices, while others believe it is simply good luck (variance of exogenous shocks hitting the U.S. economy dropping sharply). Regardless of the reason, this is an important feature of the U.S. GDP data that should be taken into account in considering the robustness of our results.

One major concern with not addressing the Great Moderation is that the linear models would be at a great disadvantage in our analysis because linear models cannot "automatically" pick up a reduction in variance, while the bounceback model can potentially proxy for the structural break in variance or other forms of heteroskedasticity through its Markov-switching structure. So the superior performance of the bounceback model may be due to capturing the break in variance rather than the asymmetries related to the business cycle. Therefore, to make sure that our results are robust, we consider a break in the variance of real GDP growth in 1984Q1 for all five time-series models presented earlier.

To implement the structural break, we consider non-parametric residuals for all of the linear models. This implies that the residuals or error terms for each of the time-series model are drawn with replacement from two separate groups stemming from the original estimation residuals, pre-structural break (1948Q4 to 1984Q1) and post-structural break (1984Q2 to 2007Q4), depending on the quarter being simulated. For the BBV model, we simulate data from a parametric model that allows for a structural break in the residual variance in 1984Q1.

Table 5 reports the results of the time-series models' abilities to reproduce business cycle features when taking into account the Great Moderation. Looking at the univariate models first, one can see that the basic findings are very similar to those reported in Table 3. The AR(2) model fails to reproduce the exact same features as it did before taking the structural break into account (average quarterly growth rates of Recovery phase, standard deviation of quarterly growth rates of Recession and Mature expansion phases, and correlation between cumulative growth rates of Recession and Recovery phases). The one noticeable difference is that the median value of the 10,000 simulated series for the correlation feature is now negative (–0.07), which is somewhat more compatible with the sample feature than the small positive correlation (+0.07) it generated before taking the structural break into account. However, the correlation is still much smaller in magnitude than the sample feature (–0.66) using the NBER chronology.

As for the bounceback model, there is very little change in terms of the results after the imposition of the structural break. Interestingly, though, compared to BBV model without a structural break, the BBV model with a structural break simulates a negative median value for the correlation between cumulative growth in the Recession phase and cumulative growth in the subsequent Recovery phase (–0.53) closer to that exhibited by actual real GDP growth using NBER chronology (–0.66).

The most interesting results in Table 5 relate to the multivariate models. There appears to be a dramatic improvement in the performance of all the multivariate models, especially the KPSW VECM. The models are now better at matching the variation in the quarterly growth rates of business cycle phases. But perhaps the most notable change is in the correlation feature. The multivariate models are now able to generate a more negative correlation between the cumulative growth rates of Recession and Recovery phases such that the proportion of simulated data below the corresponding NBER sample feature value is just slightly above 10%. This result is quite surprising given that none of the univariate linear specifications in Morley and Piger (2006) report a proportion higher than 10%. Even some univariate nonlinear models in Morley and Piger (2006) report percentiles that are far less than 10%.

However, one should be cautious in interpreting this result as a validation of multivariate linear models in terms of their ability to capture business cycle asymmetries exhibited by real GDP. First of all, the median correlations for the 10,000 simulations for all the multivariate linear models are still only mildly negative. The B&Q model generates the most negative median correlation at –0.24, which is much closer to zero than that reported for the BBV model (–0.53) and the sample correlation (–0.66). Furthermore, the fact that the multivariate linear models cannot produce a strong enough negative correlation before taking into account the structural break in variance implies there is something about the volatility reduction in 1984 that helped generate it, rather than something inherent in the dynamics of the linear models.

5.6. Counterfactual and Asymptotic Simulation Experiments

To investigate our conjecture that the stronger negative correlation between the cumulative growth rates of the Recession and the Recovery phases for the multivariate models is driven by the one-time structural break in GDP variance, we conduct two experiments involving constructing counterfactuals and using an asymptotic simulation.

If there is something about the linear dynamics in the multivariate models that allow them to capture the strong negative correlation between growth in recessions and growth in recoveries exhibited by real GDP, it should be a recurring feature of the simulated data prior to the structural break date of 1984Q1 and after it as well. So we consider the following thought experiment: What would happen if the pre-1984Q1 parameters for the multivariate linear models were applied for the whole sample period? Would this generate a strong enough negative correlation between the growth rates in recessions and recoveries? Similarly, what would the correlation be if the post-1984Q1 parameters for the multivariate linear models were applied for the whole sample period?

These questions lead us to a simple counterfactual experiment where we estimate each of the multivariate models using pre-1984Q1 data and post-1984Q1 data separately. We then assume that the pre (post) break date parameters apply to the whole sample period and simulate corresponding counterfactual data to calculate the implied correlation between the cumulative growth rate of the Recession phase and the Recovery phase. We consider both parametric and non-parametric residual specifications, although the results are very similar. Table 6 details the outcome of the counterfactual experiment.

It is clear from the table that a strong negative correlation between growth rates in recession and recovery phases is not a recurring feature using either pre or post break date parameters for any of the multivariate linear models. Under counterfactual 1 (pre-1984Q1 parameters), the median correlations for the simulations are only slightly negative or zero. With low corresponding percentiles, these results show that it is very unlikely that the sample value could have arisen from such models. Under counterfactual 2 (post-1984Q1 parameters), the median correlations for the simulations for all of the multivariate linear models are actually positive, although the corresponding percentiles are within the 0.1 to 0.9 range.

To further investigate the negative correlation feature for the multivariate linear models, we also conducted an "asymptotic" simulation exercise. If the strong negative correlations produced by the multivariate linear models are driven by the one-time structural break in variance, we should see the effect of the structural break weaken as we increase the sample size for the simulated data. Table 7 reports the correlation between the cumulative growth in Recession phase and the cumulative growth in Recovery phase for the bounceback model as well as the three multivariate linear models for an extended simulation sample period from 1884Q1 to 2084Q1 (100 years before the structural break date of 1984Q1 to 100 years after). Results show that even though the median simulated correlation remains negative for the multivariate linear models, the proportion of the 10,000 simulated features falling below that reported for the actual real GDP growth data (-0.66) is now close to zero. For the bounceback model, the median correlation remains negative, and the percentile stays above the 10% cutoff point.

Based on these two experiments, we have found some evidence to support our conjecture that the multivariate linear models with break in variance in 1984Q1 are not really capturing the negative correlation between the cumulative growth rates of the Recession and Recovery phases, but rather reflect the effect of a one-time structural break. Meanwhile, even if we take the results reported in Table 5 at face value, compared to the preferred model before imposing the structural break (BBV), the best performing linear model in Table 5 (KPSW) still fares worse in terms of reproducing business cycle features. Most importantly, these results illustrate that, while a more general model will always fit the data better in sample, it does not necessarily do better in other dimensions.[10]

6. Conclusion

In this paper, we assessed the ability of various time-series models to reproduce business cycle features exhibited by U.S. real GDP. Following Morley and Piger (2006), we use an accurate business cycle dating algorithm to calculate business cycle turning points for the simulated data from each of the time-series models. The univariate linear and nonlinear models and the multivariate linear models we consider here allow us to answer the question of whether multivariate information can enrich the linear models such that they would succeed where univariate linear models have failed in terms of replication of certain business cycle features.

From the simulation exercises, a few important results emerge. First of all, the use of a semi-parametric bootstrap approach to residual specification seems to benefit some models, particularly the linear models. At the same time, the fact that the linear models with non-parametric residuals fail to capture the strong negative correlation between the cumulative growth of the Recession phase and the cumulative growth of the Recovery phase, while the bounceback model with normal parametric residuals does capture this feature, suggests that the failure of the linear models is not due to misspecification of the error terms. Perhaps the semi-parametric bootstrap procedure improved the performance of the linear models only because it

[10] This is analogous to the idea that a more parsimonious model can forecast better out-of-sample, even if it fits worse in sample.

allowed the linear models to better capture fat tails or skewness present in the unconditional distribution of real GDP growth, something that the parametric nonlinear models already take into account.

Secondly, the imposition of a structural break in the variance of real GDP growth in 1984Q1 had a noticeable impact on the performance of the multivariate linear models, enabling the VAR and VECM models to come closer to matching the BBV model's ability to replicate most of the business cycle features considered here. However, our counterfactual and asymptotic simulation experiments show that this improvement is not due to an ability to produce recurring patterns in the data, but merely reflects the one-time structural break.

Finally, the nonlinear bounceback model is by far the best performing time-series model among the ones we consider here. It can capture not only the usual features other papers in the literature report, such as the length and variation of business cycle phases or the average and standard deviation of quarterly growth rates of business cycle phases, but also important business cycle asymmetries that economists have observed in the GDP data. Specifically, the bounceback model succeeds at replicating the higher than average growth rates during the Recovery phase and the strong correlation between the severity of a recession and the strength of the subsequent recovery. This result is consistent with findings in Morley and Piger (2006) and corroborates the results in Galvão (2003) and Demers and Macdonald (2007). What this suggests is that there is nonlinearity present in the U.S. business cycle that cannot be picked up just by introducing variables such as the unemployment rate, inflation, interest rates, and the components of GDP into linear models. Instead, there is something fundamentally different about the dynamics of real GDP coming out of a recession that linear models simply are not able to replicate.

Appendix

Here we present the estimates for quarterly U.S. GDP for the five time-series models under consideration. The reported estimates are used to calibrate the data generating process used in our Monte Carlo simulations. The AR(2) and the Kim, et. al. (2005) bounceback model are univariate, while the Blanchard and Quah (1989) VAR, the Ahmed et. al. (2004) VAR, and the King et. al. (1991) VECM are multivariate. For the univariate models, Δy_t is defined as annualized growth rate of output to be compatible with the specification in Morley and Piger (2006). For the multivariate models, Δy_t is defined as natural log difference of output to be compatible with their original specifications.

<u>The AR(2) model:</u>

Estimation period 1948Q4 to 2007Q4.

$$\Delta y_t = 0.0214 + 0.2976\Delta y_{t-1} + 0.0858\Delta y_{t-2} + \varepsilon_t ,$$

$$\sigma_\varepsilon = 0.0383 .$$

<u>The Kim, Morley, and Piger (2005) bounceback model (BBV):</u>

Estimation period 1948Q4 to 2007Q4.

$$\Delta y_t = 3.3521 - 4.4383 S_t + 1.3052(1 - S_t)\sum_{j=1}^{6} S_{t-j} + \varepsilon_t ,$$

$$\sigma_\varepsilon = 3.1122 ,\ P(S_t = 1 \mid S_{t-1} = 1) = 0.7321 ,\ P(S_t = 0 \mid S_{t-1} = 0) = 0.9450 ,$$

where $S_t = 1$ corresponds to recessions and $S_t = 0$ corresponds to expansions.

<u>The Kim, Morley, and Piger (2005) bounceback model with break in variance (BBV):</u>

Estimation period 1948Q4 to 2007Q4.

$$\Delta y_t = 3.1464 - 4.4459 S_t + 1.5110(1 - S_t)\sum_{j=1}^{6} S_{t-j} + \varepsilon_t ,$$

$\sigma_\varepsilon = 4.0732$ for t = 1948Q4 to 1984Q1,

$\sigma_\varepsilon = 1.9881$ for t = 1984Q2 to 2007Q4,

$P(S_t = 1 | S_{t-1} = 1) = 0.7630$, $P(S_t = 0 | S_{t-1} = 0) = 0.9716$,

where $S_t = 1$ corresponds to recessions and $S_t = 0$ corresponds to expansions.

Blanchard & Quah (1989) two-variable VAR model (B&Q):

Estimation period 1950Q1 to 2007Q4.

$$\begin{aligned}\Delta y_t &= 0.0022 + 0.1254\Delta y_{t-1} + 0.1682\Delta y_{t-2} + 0.0532\Delta y_{t-3} + 0.1426\Delta y_{t-4} + 0.06208\Delta y_{t-5} \\ &\quad + 0.1596\Delta y_{t-6} - 0.0158\Delta y_{t-7} + 0.0231\Delta y_{t-8} - 0.7470u_{t-1} + 1.5542u_{t-2} - 0.5442u_{t-3} \\ &\quad + 0.5880u_{t-4} - 0.8945u_{t-5} + 0.3827u_{t-6} - 0.2552u_{t-7} - 0.0012u_{t-8} + \varepsilon_t,\end{aligned}$$

$$\Sigma_\varepsilon = \begin{bmatrix} 0.0000762302 & -0.0000148006 \\ -0.0000148006 & 0.0000070473 \end{bmatrix},$$

where u_t denotes the civilian unemployment rate and the order of the variables in the VAR is $[\Delta y_t \ u_t]'$. The quarterly unemployment rate is the average of the monthly unemployment rate series.

Ahmed, Levin, and Wilson (2004) four-variable VAR model (ALW):

Estimation period 1955Q3 to 2007Q4.

$$\begin{aligned}\Delta y_t &= 0.0076 + 0.2145\Delta y_{t-1} + 0.1660\Delta y_{t-2} + 0.0021\Delta y_{t-3} - 0.0328\Delta y_{t-4} + 0.0673\Delta cpi_{t-1} \\ &\quad - 0.0316\Delta cpi_{t-2} + 0.0214\Delta cpi_{t-3} - 0.1648\Delta cpi_{t-4} - 0.0144\Delta ppi_{t-1} + 0.0263\Delta ppi_{t-2} \\ &\quad - 0.0238\Delta ppi_{t-3} + 0.0139\Delta ppi_{t-4} + 0.0160 ffr_{t-1} - 0.2538 ffr_{t-2} + 0.1140 ffr_{t-3} + 0.0998 ffr_{t-4} \\ &\quad + \varepsilon_t,\end{aligned}$$

$$\Sigma_\varepsilon = \begin{bmatrix} 0.000066 & -0.000001 & 0.000006 & 0.000023 \\ -0.000001 & 0.000021 & 0.000039 & 0.000012 \\ 0.000006 & 0.000039 & 0.000165 & 0.000039 \\ 0.000023 & 0.000012 & 0.000039 & 0.000120 \end{bmatrix},$$

where Δcpi_t denotes the consumer price inflation rate, Δppi_t is the inflation rate of the producer price index: all commodities, and ffr_t is the federal funds rate. The order of the variables in the VAR is $[\Delta y_t \ \Delta cpi_t \ \Delta ppi_t \ ffr_t]'$. The quarterly *cpi*, *ppi*, and *ffr* are all constructed by picking the end of quarter value of the equivalent monthly series.

King, Plosser, Stock, and Watson (1991) three-variable VECM (KPSW):

Estimation period 1949Q2 to 2007Q4.

$$
\begin{aligned}
\Delta y_t &= 0.0008 + 0.0895\left(c_{t-1} - y_{t-1} + 0.4178\right) - 0.0265(i_{t-1} - y_{t-1} + 2.0545) + 0.1462\Delta y_{t-1} \\
&\quad + 0.0526\Delta y_{t-2} + 0.0276\Delta y_{t-3} - 0.0636\Delta y_{t-4} + 0.1650\Delta y_{t-5} + 0.0752\Delta y_{t-6} - 0.0865\Delta y_{t-7} \\
&\quad + 0.0057\Delta y_{t-8} + 0.2790\Delta c_{t-1} + 0.1360\Delta c_{t-2} - 0.0079\Delta c_{t-3} + 0.1432\Delta c_{t-4} - 0.1311\Delta c_{t-5} \\
&\quad - 0.0009\Delta c_{t-6} + 0.1878\Delta c_{t-7} - 0.0724\Delta c_{t-8} + 0.0134\Delta i_{t-1} + 0.0202\Delta i_{t-2} - 0.0084\Delta i_{t-3} \\
&\quad + 0.0138\Delta i_{t-4} - 0.0333\Delta i_{t-5} - 0.0026\Delta i_{t-6} + 0.0121\Delta i_{t-7} + 0.0100\Delta i_{t-8} + \varepsilon_t,
\end{aligned}
$$

$$
\Sigma_\varepsilon = \begin{bmatrix} 0.000078 & 0.000039 & 0.000270 \\ 0.000039 & 0.000057 & 0.000053 \\ 0.000270 & 0.000053 & 0.001631 \end{bmatrix},
$$

where c_t denotes real personal consumption expenditure and i_t is the real gross private domestic investment. The order of the variables in the VECM is $[y_t \;\; c_t \;\; i_t]'$ and the two cointegrating relationships based on the balance growth theory are $(c_t - y_t)$ and $(i_t - y_t)$.

References

Ahmed, Shaghil, Andrew Levin, and Beth Anne Wilson. 2004. "Recent U.S. Macroeconomic Stability: Good Policies, Good Practices, or Good Luck?" *The Review of Economics and Statistics*. 86(3): 824-832.

Blanchard, Olivier J., and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*. 79(4): 655-673.

Bry, Gerhard, and Charlotte Boschan. 1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. New York and London: Columbia University Press.

Boivin, Jean, and Marc P. Giannoni. 2006. "Has Monetary Policy Become More Effective?" *The Review of Economics and Statistics*, 88(3): 445-462.

Boldin, Michael D.. 1996. "A Check on the Robustness of Hamilton's Markov Switching Model Approach to the Economic Analysis of the Business Cycle." *Studies in Nonlinear Dynamics and Econometrics*. 1(1): 35-46.

Burns, Arthur F., and Wesley C. Mitchell. 1947. *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Clements, Michael P., and Hans-Martin Krolzig. 2004. "Can Regime-Switching Models Reproduce the Business Cycle Features of US Aggregate Consumption, Investment and Output?" *International Journal of Finance and Economics*, 9(1): 1-14.

Demers, Frédérick, and Ryan Macdonald. 2007. "The Canadian Business Cycle: A Comparison of Models." Bank of Canada Working Paper 2007-38.

Friedman, Milton. 1964. "Monetary Studies of the National Bureau." In *The National Bureau Enters Its 45th Year, 44th Annual Report*. New York: National Bureau of Economic Research. [Reprinted in Friedman, Milton. 1969. *The Optimum Quantity of Money and Other Essays*. Chicago: Aldine.]

Friedman, Milton. 1993. "The "Plucking Model" of Business Fluctuations Revisited." *Economic Inquiry*. 31(2): 171-177.

Hamilton, James D.. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica*. 57(2): 357-384.

Harding, Don, and Adrian Pagan. 2002. "Dissecting the Cycle: a Methodological Investigation." *Journal of Monetary Economics*. 49(2): 365-381.

Hess, Gregory D., Shigeru Iwata. 1997. "Measuring and Comparing Business-Cycle Features." *Journal of Business & Economic Statistics*. 15(4): 432-444.

Galvão, Ana Beatriz C.. 2002. “Can Non-linear Time Series Models Generate US Business Cycle Asymmetric Shape?” *Economics Letters*. 77(2): 187-194.

Kim, Chang-Jin, James Morley, and Jeremy Piger. 2005. “Nonlinearity and the Permanent Effects of Regressions.” *Journal of Applied Econometrics*. 20(2): 291-309.

King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson. 1991. “Stochastic Trends and Economic Fluctuations.” *American Economic Review*. 81(4): 819-840.

McConnell, Margaret M., and Gabriel Perez-Quiros. 2000. “Output Fluctuations in the United States: What Has Changed Since the Early 1980’s?” *American Economic Review*. 90(5): 1464-1476.

McQueen, Grant, and Steven Thorley. 1993. “Asymmetric Business Cycle Turning Points.” *Journal of Monetary Economics*. 31(3): 341-362.

Morley, James, and Jeremy Piger. 2006. “The Importance of Nonlinearity in Reproducing Business Cycle Features.” *Contributions to Economic Analysis Series*, Vol. 276, *Nonlinear Time Series Analysis of Business Cycles*, 75-94. Amsterdam and San Diego: Elsevier.

Morley, James, and Jeremy Piger. 2009. “The Asymmetric Business Cycle.” http://artsci.wustl.edu/~morley/abc.pdf.

Stock, James H., and Mark W. Watson. 2002. “Has the Business Cycle Changed and Why?” In *NBER Macroeconomics Annual*, Vol. 17, ed. Mark Gertler and Kenneth Rogoff, 159-218. Cambridge and London: MIT Press.

TABLE 1

PEAK AND TROUGH DATES FROM NBER BUSINESS CYCLE DATING COMMITTEE AND THE BBQ AND MBBQ ALGORITHMS APPLIED TO U.S. REAL GDP (1948Q4 – 2007Q4)

Business Cycle Peaks			Business Cycle Troughs		
NBER	BBQ	MBBQ	NBER	BBQ	MBBQ
1948Q4	-	-	1949Q4	**1949Q2**	1949Q4
1953Q2	1953Q2	1953Q2	1954Q2	**1954Q1**	1954Q2
1957Q3	1957Q3	1957Q3	1958Q2	**1958Q1**	**1958Q1**
1960Q2	**1960Q1**	**1960Q1**	1961Q1	**1960Q4**	**1960Q4**
1969Q4	**1969Q3**	**1969Q3**	1970Q4	1970Q4	1970Q4
1973Q4	1973Q4	1973Q4	1975Q1	1975Q1	1975Q1
1980Q1	1980Q1	1980Q1	1980Q3	1980Q3	1980Q3
1981Q3	1981Q3	1981Q3	1982Q4	**1982Q1**	1982Q4
1990Q3	1990Q3	1990Q3	1991Q1	1991Q1	1991Q1
2001Q1	-	-	2001Q4	-	-

Note: Bold indicate that the identified turning points differ from the NBER dates. We ignore the first NBER peak date in our evaluation of the BBQ and MBBQ algorithm because given our sample period, the earliest date at which the algorithms can identify a turning point is 1949Q2.

TABLE 2

BUSINESS CYCLE FEATURES FOR U.S. REAL GDP (1948Q4 – 2007Q4)

	NBER	BBQ	MBBQ
Average quarterly growth rates			
Recession	–1.92	**–2.96**	–2.49
Expansion	4.59	4.78	**4.98**
Recovery	7.10	**5.52**	7.23
Mature expansion	3.94	**4.57**	4.29
Std. deviation of quarterly growth rates			
Recession	3.33	**3.10**	3.13
Expansion	3.54	**3.83**	3.75
Recovery	4.18	**4.75**	4.25
Mature expansion	3.05	**3.51**	3.31
Number of phases			
Number of peaks	9	8	8
Average length of phases			
Recession	3.44	**3.00**	3.50
Expansion	19.67	17.88	**17.13**
Std. deviation of length of phases			
Recession	1.13	1.31	**1.41**
Expansion	12.72	11.34	**10.88**
Correlation between growth rates			
Recession/Recovery	–0.66	**–0.36**	–0.68

Note: Because the earliest date at which the algorithms can identify a turning point is 1949Q2, we ignore the first peak in 1948Q4 when calculating the sample features associated with the NBER dates. Bold indicates that the feature values produced by the algorithm is "further away" from the NBER sample feature values.

TABLE 3

BUSINESS CYCLE FEATURES FOR UNIVARIATE MODELS
(1948Q4 – 2007Q4)

Features	Real GDP	AR(2) (parametric residuals)	AR(2) (non-parametric residuals)	BBV (parametric residuals)	BBV (non-parametric residuals)
Average quarterly growth rates					
Recession	-1.92	-2.06 (0.63)	-2.19 (0.71)	-2.12 (0.69)	**-2.67 (0.93)**
Expansion	4.59	4.29 (0.80)	4.11 (0.89)	4.19 (0.89)	**4.16 (0.90)**
Recovery	7.10	**4.16 (1.00)**	**3.98 (1.00)**	5.87 (0.90)	6.16 (0.83)
Mature expansion	3.94	4.31 (0.18)	4.13 (0.33)	3.83 (0.66)	3.74 (0.76)
Std. deviation of quarterly growth rates					
Recession	3.33	**2.27 (0.99)**	**2.54 (0.96)**	**2.34 (0.98)**	2.81 (0.86)
Expansion	3.54	3.56 (0.46)	3.59 (0.43)	3.56 (0.46)	3.71 (0.27)
Recovery	4.18	**3.23 (0.97)**	3.19 (0.87)	4.02 (0.61)	4.09 (0.56)
Mature expansion	3.05	**3.62 (0.01)**	**3.63 (0.04)**	3.33 (0.13)	**3.45 (0.09)**
Number of phases					
Number of peaks	9	9 (0.40)	8 (0.61)	9 (0.50)	8 (0.55)
Average length of phases					
Recession	3.44	3.27 (0.60)	3.29 (0.60)	3.45 (0.49)	3.75 (0.36)
Expansion	19.67	21.00 (0.42)	24.43 (0.24)	22.33 (0.33)	22.88 (0.31)
Std. deviation of length of phases					
Recession	1.13	1.56 (0.27)	1.51 (0.29)	1.83 (0.20)	2.07 (0.14)
Expansion	12.72	16.30 (0.27)	19.08 (0.17)	17.14 (0.24)	17.70 (0.21)
Correlation between growth rates					
Recession/Recovery	-0.66	**0.07 (0.04)**	**0.06 (0.06)**	-0.44 (0.24)	-0.49 (0.29)

Note: First column reports the U.S. real GDP growth sample features using NBER peak and trough dates. Following columns report simulated median feature for the univariate models based on 10,000 simulations, with the proportion of simulated features that fall below the sample feature reported in column 1 in brackets. Bold indicates a percentile that is less than 0.1 or greater than 0.9, implying that it was unlikely that the particular time-series model could simulate data that replicates the behavior of actual GDP for that particular feature.

TABLE 4

BUSINESS CYCLE FEATURES FOR MULTIVARIATE MODELS
(1948Q4 – 2007Q4)

Features	Real GDP	B&Q (parametric residuals)	B&Q (non-parametric residuals)	ALW (parametric residuals)	ALW (non-parametric residuals)	KPSW (parametric residuals)	KPSW (non-parametric residuals)
Average quarterly growth rates							
Recession	-1.92	-2.07 (0.65)	-2.12 (0.68)	-1.85 (0.42)	-1.88 (0.46)	-2.13 (0.71)	-2.14 (0.68)
Expansion	4.59	4.35 (0.78)	4.22 (0.90)	**3.97 (0.97)**	**3.81 (0.99)**	4.24 (0.84)	**4.01 (0.95)**
Recovery	7.10	**4.71 (1.00)**	**4.69 (1.00)**	**4.11 (1.00)**	**3.96 (1.00)**	**4.59 (1.00)**	**4.33 (1.00)**
Mature expansion	3.94	4.26 (0.17)	4.10 (0.31)	3.93 (0.51)	3.77 (0.69)	4.13 (0.32)	3.92 (0.53)
Std. deviation of quarterly growth rates							
Recession	3.33	**2.13 (1.00)**	**2.25 (1.00)**	**2.01 (1.00)**	**2.21 (0.97)**	**2.25 (1.00)**	**2.45 (0.95)**
Expansion	3.54	3.66 (0.28)	3.55 (0.48)	3.41 (0.73)	3.29 (0.80)	3.70 (0.23)	3.49 (0.56)
Recovery	4.18	**3.43 (0.94)**	3.36 (0.89)	**3.14 (0.98)**	**2.98 (0.94)**	**3.47 (0.93)**	**3.26 (0.93)**
Mature expansion	3.05	**3.70 (0.00)**	**3.56 (0.04)**	**3.46 (0.03)**	3.31 (0.19)	**3.74 (0.00)**	**3.52 (0.05)**
Number of phases							
Number of peaks	9	10 (0.24)	9 (0.40)	9 (0.34)	8 (0.59)	11 (0.14)	9 (0.37)
Average length of phases							
Recession	3.44	3.13 (0.72)	3.00 (0.80)	3.11 (0.72)	3.00 (0.77)	3.30 (0.60)	3.20 (0.67)
Expansion	19.67	19.10 (0.55)	21.63 (0.36)	20.22 (0.46)	24.43 (0.24)	17.18 (0.70)	20.89 (0.42)
Std. deviation of length of phases							
Recession	1.13	1.26 (0.39)	1.15 (0.48)	1.32 (0.37)	1.21 (0.45)	1.41 (0.27)	1.30 (0.36)
Expansion	12.72	13.59 (0.43)	15.89 (0.30)	15.43 (0.31)	18.85 (0.19)	12.08 (0.55)	14.94 (0.35)
Correlation between growth rates							
Recession/Recovery	-0.66	**-0.07 (0.04)**	**-0.14 (0.08)**	**0.01 (0.04)**	**-0.01 (0.07)**	**-0.09 (0.04)**	**-0.12 (0.07)**

Note: First column reports the U.S. real GDP growth sample features using NBER peak and trough dates. Following columns report simulated median feature for the multivariate models based on 10,000 simulations, with the proportion of simulated features that fall below the sample feature reported in column 1 in brackets. Bold indicates a percentile that is less than 0.1 or greater than 0.9, implying that it was unlikely that the particular time-series model could simulate data that replicates the behavior of actual GDP for that particular feature.

TABLE 5

BUSINESS CYCLE FEATURES FOR ALL MODELS WITH STRUCTURAL BREAK
(1948Q4 – 2007Q4 WITH STRUCTURAL BREAK IN VARIANCE IN 1984Q1)

Features	Real GDP	AR(2)	BBV	B&Q	ALW	KPSW
Average quarterly growth rates						
Recession	-1.92	-2.36 (0.79)	-2.18 (0.69)	-2.18 (0.71)	-2.10 (0.63)	-2.29 (0.76)
Expansion	4.59	4.31 (0.70)	**3.94 (0.96)**	4.29 (0.81)	**3.94 (0.95)**	4.24 (0.80)
Recovery	7.10	**4.55 (0.99)**	5.69 (0.88)	**5.00 (0.98)**	**4.33 (1.00)**	**4.82 (0.99)**
Mature expansion	3.94	4.26 (0.27)	3.62 (0.86)	4.13 (0.29)	3.84 (0.60)	4.10 (0.35)
Std. deviation of quarterly growth rates						
Recession	3.33	**2.67 (0.90)**	**2.39 (0.95)**	**2.33 (0.99)**	**2.47 (0.92)**	2.61 (0.89)
Expansion	3.54	3.81 (0.25)	3.66 (0.34)	3.70 (0.31)	3.47 (0.58)	3.64 (0.38)
Recovery	4.18	3.76 (0.69)	4.30 (0.44)	3.75 (0.72)	3.41 (0.83)	3.70 (0.76)
Mature expansion	3.05	**3.79 (0.05)**	3.42 (0.11)	**3.63 (0.04)**	3.43 (0.14)	**3.59 (0.06)**
Number of phases						
Number of peaks	9	7 (0.70)	7 (0.72)	8 (0.52)	8 (0.62)	8 (0.57)
Average length of phases						
Recession	3.44	3.33 (0.56)	3.43 (0.51)	3.00 (0.77)	3.10 (0.73)	3.20 (0.67)
Expansion	19.67	22.22 (0.36)	26.00 (0.19)	22.00 (0.34)	22.57 (0.34)	21.88 (0.37)
Std. deviation of length of phases						
Recession	1.13	1.60 (0.28)	1.83 (0.24)	1.17 (0.46)	1.28 (0.40)	1.30 (0.37)
Expansion	12.72	18.89 (0.23)	20.27 (0.15)	17.28 (0.23)	18.21 (0.22)	17.67 (0.25)
Correlation between growth rates						
Recession/Recovery	-0.66	**-0.07 (0.10)**	-0.53 (0.35)	-0.24 (0.13)	**-0.13 (0.10)**	-0.20 (0.12)

Note: First column reports the U.S. real GDP growth sample features using NBER peak and trough dates. Following columns report simulated median feature for all the time-series models based on 10,000 simulations, with the proportion of simulated features that fall below the sample feature reported in column 1 in brackets. Bold indicates a percentile that is less than 0.1 or greater than 0.9, implying that it was unlikely that the particular time-series model could simulate data that replicates the behavior of actual GDP for that particular feature.

TABLE 6

COUNTERFACTUAL EXPERIMENT RESULT FOR MULTIVARIATE MODELS

Correlation between Cumulative Growth in Recession Phase and Cumulative Growth in Recovery Phase	Pre-structural break Parameters (Counterfactual 1)	Post-structural break Parameters (Counterfactual 2)
Real GDP	-0.66	-0.66
B&Q		
Parametric	**-0.11 (0.03)**	0.00 (0.28)
Non-parametric	**-0.14 (0.04)**	0.00 (0.34)
ALW		
Parametric	**0.00 (0.02)**	0.04 (0.27)
Non-parametric	**0.00 (0.02)**	0.00 (0.32)
KPSW		
Parametric	**-0.12 (0.03)**	0.14 (0.18)
Non-parametric	**-0.15 (0.05)**	0.09 (0.27)

Note: First row reports the U.S. real GDP growth sample features using NBER peak and trough dates. Following rows report simulated median feature for the multivariate models based on 10,000 simulations, with the proportion of simulated features that fall below the sample feature reported in row 1 in brackets. Bold indicates a percentile that is less than 0.1 or greater than 0.9, implying that it was unlikely that the particular time-series model could simulate data that replicates the behavior of actual GDP for that particular feature. The structural break date is 1984Q1.

TABLE 7

ASYMPTOTIC SIMULATION EXPERIMENT RESULT

	Correlation between Cumulative Growth in Recession Phase and Cumulative Growth in Recovery Phase
Real GDP	-0.66
BBV	-0.46 (0.11)
B&Q	**-0.26 (0.01)**
ALW	**-0.16 (0.01)**
KPSW	**-0.24 (0.01)**

Note: First row reports the U.S. real GDP growth sample features using NBER peak and trough dates. Following rows report simulated median feature for the bounceback and multivariate linear models based on 10,000 simulations of length 200 years, with the proportion of simulated features that fall below the sample feature reported in row 1 in brackets. Bold indicates a percentile that is less than 0.1 or greater than 0.9, implying that it was unlikely that the particular time-series model could simulate data that replicates the behavior of actual GDP for that particular feature. The structural break date is 1984Q1.