Essays on the Impacts of the Great Moderation on Business Cycle Modeling

by

Andre R. Neveu

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York-Graduate Center

2009

This manuscript has been read and accepted for the
Graduate Faculty in Economics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy

Dr. Merih Uctum

Date

Chair of Examining Committee

Dr. Thom Thurston

Date

Executive Officer

Dr. Merih Uctum

Dr. Sangeeta Pratap

Dr. Temisan Agbeyegbe

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

ESSAYS ON THE IMPACTS OF THE GREAT MODERATION ON BUSINESS CYCLE MODELING

by

Andre R. Neveu

Adviser: Professor Merih Uctum

The research presented here is a comprehensive analysis of research on the "Great Moderation" and its impact on business cycle modeling. In the presence of a less volatile aggregate economy, the methods of modeling business cycles have fundamentally changed along with the ability to detect turning points in the business cycle using standard algorithms. Chapter One lays out the historical case for modeling the business cycle in a manner placing importance on the ability of a model to replicate features observed in actual GDP data, such as the depth and length of recessions, or the average height of expansions. Chapter Two compares different business cycle models by their ability and accuracy in reproducing features of observed GDP data in simulated Monte Carlo paths. Comparisons are made by examining how volatility moderation affects business cycle modeling for the U.S., U.K., and Australia. Univariate ARIMA, structural change, and Markov-switching ("MS") models are estimated and used to simulate time paths using Monte Carlo methods. These results generally support previous findings that MS models are superior to linear models and comparable to structural change models at fitting business cycle characteristics. Tests show that to replicate business cycle characteristics, MS models must account for independent shifts in mean and volatility parameters. Substantial new evidence shows that commonly specified MS models with a simple linear structure, constant variance, or state-dependent volatility are sub-optimal and should be avoided in practice. Results indicate that models attempting to replicate business cycle features in any series should consider the importance of how volatility is modeled prior to estimation. Evidence is also presented showing that the Great Moderation may have recently ended. Chapter Three examines algorithm robustness used to conclude that independent switching models are better able to replicate business cycle features. Robustness is tested by varying the parameters of dating algorithms used to detect turning points. Evidence shows that the “window” and “censor” used for turning point selection criteria does not lead to substantial changes in the conclusions of our previous findings implying that our results are not artifacts of the algorithm, but due to the actual economic model itself.

Contents

List of Tables

List of Figures

Chapter 1

A Review of Business Cycle Modeling and the Great Moderation

1.1 Introduction

Macroeconomic business cycle models are often used to make statements about current and future cyclical conditions of an economy. Predictions about the economy are often made using models in the context of a set of stylized facts that have been compiled using historical data. Typically, these predictions are made either knowing or ignoring the fact that the predictive model might not be able to simulate data that statistically resembles the cyclical features of interest. Furthermore, the multitude of models that a researcher can choose from could lead to confusing or conflicting conclusions about how long or deep a cycle might last.

Adelman and Adelman (1959) were among the first to attempt to test the dynamics of a macroeconomic model for the ability to reproduce stylized facts of aggregate data. Following in the spirit of Frisch (1933), Adelman and Adelman tested a large Klein-Goldberger model for dynamic behavior by using random perturbations to see if simulated time paths exhibited similar business cycle fluctuations relative to those seen in aggregate data. Since 1959, many other researchers have subsequently tested macroeconomic models for their ability to replicate business cycle features using perturbations and Monte Carlo methods in what is often referred to as the "Test of the Adelmans."

The research presented here employs testing methods developed in the spirit of the Adelmans to examine some commonly used business cycle models in an effort to help improve the model selection process. In particular, this research expands the model selection process to incorporate the fact that the mid-1980's to 2008 experienced much lower volatility in many aggregate variables relative to the post-war period. Examining the U.S., U.K., and Australia it is concluded that the less volatile period, known as the "Great Moderation" must be properly accounted for in a business cycle model in order to reproduce the stylized facts of the observed business cycle. The stylized facts that are measured and compared here include features such as the length or depth of recessions, and the height and accumulated output from expansions. If models are unable to match these features, it is likely that the predictions are inaccurate or misleading.

Chapter Two shows that commonly used business cycle models, including linear, structural change in volatility ("structural change" hereafter), and Markov-switching ("MS" hereafter) approaches should examine data created through simulation to compare the features of the simulation to those seen in the underlying series. Standard statistical tests do not apply to the various MS model specifications used here. Therefore standard tests for parametric fit and model specification are unable to select models that are able to match cyclical features with confidence. Even though MS models are intended to help select turning points parametrically, many of the models frequently used in the literature are unable to simulate data that resembles features seen in actual Gross Domestic Product ("GDP" hereafter) data. The preferred specification for the three countries studied here shows new evidence that the Great Moderation may be ending. Chapter Three presents evidence that this model selection process is robust to the parameters of the algorithms used to measure business cycle features in the original and simulated data. As long as a consistent algorithm is used to measure actual and simulated data the same models tend to be selected. Our preference for model selection here is based on the ability of a model to replicate actual business cycle features in simulated data.

1.2 A History of Business Cycle Dating

The field of business cycle modeling can be traced back to the work of Burns and Mitchell (1946) who were originally seeking to identify turning points in a variety of macroeconomic time series to make statements about the general cyclical behavior of the overall economy. By locating the turning points in several different aggregate variables, Burns and Mitchell measured the *specific cycles*, or cyclical behavior and features of these series. Authors such as Burns and Mitchell were also interested in understanding the relation between turning points in specific cycles relative to the chronology of turning points in general aggregate business activity. The *reference cycle* as used here is defined as the behavior and chronology of turning points in general aggregate activity. The definition of the reference cycle used here differs from the definition used by Burns and Mitchell (1946) which referred to the characteristics of a specific activity relative to the turning points in reference activity. We depart from this naming convention here as we are trying to estimate the cycle in the reference activity itself.

By combining the information on specific cycles in several series Burns and Mitchell gained a better understanding of the relationship between sectors of the economy and the *reference dates* of the reference cycle, which represented the peaks and troughs of the aggregate economy. This effort can be seen as an early attempt at identifying leading and lagging indicators of the cyclical status of the aggregate economy. It is important to note that there is only one set of reference dates for the economy signaling the beginning and ending dates of recessionary and expansionary periods. While individual series will show specific cycles, the aggregation of information gleaned from specific cycles leads to statements about the reference cycle.

The traditional method for dating the *classical cycle* using aggregate data can be traced back to Burns and Mitchell (1946). Bry and Boschan (1971) later examined specific cycle traits relative to the reference dates that were determined by the the National Bureau of Economic Research. Harding and Pagan (2002) adapted the methods of Bry and Boschan to help identify turning points in the reference cycle itself using only quarterly GDP data. Subsequently, many others have incorporated the methods of Harding and Pagan to study classical business cycle models. The classical cycle refers only to the cyclical nature of output and not modern or *growth cycles* which measure trends in growth.

1.2.1 The NBER-BCDC

Contrary to the beliefs of many, a recession is not defined as a period where GDP contracts for two consecutive quarters. In the U.S. the National Bureau of Economic Research ("NBER") Business Cycle Dating Committee ("BCDC") is the official determinant of business cycle turning points. The NBER-BCDC ("NBER" hereafter) defines a recession as "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales (National Bureau of Economic Research, 2003)." The NBER uses several *specific* macroeconomic series to identify a turning point in economic activity, acknowledging that they are attempting to determine only a single turning point in the business or *reference* cycle.

The NBER uses publicly available monthly data on personal income, employment, industrial production, and final sales in manufacturing and retail, to help them determine the peak and trough dates in the overall economy (National Bureau of Economic Research, 2003). Estimates for monthly GDP are acquired from the private firm, Macroeconomic Advisers, and incorporated into NBER statements. One of the primary reasons that the NBER uses multiple series to identify turning points in the aggregate economy is due to the fact that official GDP data is only available on a quarterly basis. While the NBER officially determines recession start and end dates, there is no universally transparent method that can be applied by outside observers which is able to perfectly mimic the NBER's choices.

As seen in Table 1.1, the NBER releases summarized stylized facts about business cycle durations on a monthly basis. The quarterly adjustments in Table 1.1 are calculated by dividing the monthly durations by three. Official NBER quarterly peaks and troughs are determined independently of the monthly turning points which is why in some cases in the past the two have not appeared consistent. In the post-war period, the typical recession in the U.S. has lasted approximately 3.3 quarters. The average expansion has lasted 19 quarters, or 4.75 years.

While the statistics in Table 1.1 are very useful at trying to improve upon macroeconomic models of U.S. business cycles, there is little guidance about how to transfer the NBER methods to other countries. With the inability to fully replicate the NBER methods over even a short period, researchers have turned to computational algorithms to help detect turning points in GDP or any series. Today, business cycle research often attempts to mimic the turning point decisions of the NBER. It was the NBER who recently declared a peak, or turning point, in the U.S. business cycle in December 2007. Harding and Pagan (2002) among others only use quarterly GDP information to determine reference dates, and features of the aggregate business cycle finding that the reference dates generally agree with those of the NBER.

The Centre for Economic Policy Research Business Cycle Dating Committee ("CEPR-BCDC") is a European economic research organization similar to the NBER, but their dating methods are not identical (Centre for Economic Policy Research, 2009). The CEPR-BCDC reports their turning points based on quarterly data, and uses real GDP aggregated over the Euro area as the "main measure of economic activity (Centre for Economic Policy Research, 2009)." Neither the NBER, or CEPR-BCDC make predictions about turning points, but rather serve to establish a chronology of facts. The business cycle research of Burns and Mitchell (1946) and Bry and Boschan (1971) was not primarily intended to be used for analyzing current economic conditions. However, much effort has been expended to incorporate the historical record of business cycle analysis into contemporary models of the economy. Other countries, such as the U.K. or Australia do not have official business cycle dating bodies, but many studies exist providing estimates for the turning points for these countries.

1.2.2 Filtering Data and the Classical Cycle

An important distinction between the work done here and much other business cycle dating research is that the data is never detrended using a filter. All data is detrended using the first difference of the log-level of GDP to make the data stationary. As noted by Harding and Pagan (2002), Burns and Mitchell did not prefer to deal with detrended data when trying to identify turning points in the classical cycle. Commonly, researchers have used filters such as the Hodrick-Prescott, Baxter-King, or band-pass to separate the growth cycle from the business cycle. However, this practice has proved troublesome in the past, as research by Canova (1994), Canova (1998), and Cogley and Nason (1995) debate the merits of applying filters to extract only the cyclical data from a time series. In the case of the Hodrick-Prescott and band-pass filters, filtered data can exhibit cyclical behavior when no cycles exist in the data. Harding and Pagan (2002), "...illustrate that trend and cycle are inextricably entwined, it would be better if we described such filters as removing the permanent component, as it is certainly untrue that their residual is the business cycle." What is clear from the literature on whether or not one should filter data, is that the cyclical characteristics that remain after detrending are specific to the particular type of filter that is used.

In Burns and Mitchell (1946) the relationship between secular and cyclical behavior of a series is discussed in detail, concluding that, "...[i]t is fairly common for statisticians to assume that the elimination of the secular trend from a time series indicates what the course of the series would have been in the absence of secular movements, and that the graduation of a time series, whether in original or trend-adjusted form, indicates what the course of the series would have been in the absence of random movements. There is no warrant for such simple interpretations (Burns and Mitchell, 1946, p. 38)." Burns and Mitchell also note, "It may be legitimate for students concerned with secular trends to put cyclical fluctuations out of sight, but students of cyclical behavior cannot take similar liberty with secular trends."

Recent research by Murray (2003) reinforces the point made by Harding and Pagan (2002) and Burns and Mitchell with more up-to-date filtering methods, showing that neither band pass filtering nor the Baxter-King filter isolate cyclical features, but "...the properties of the filtered series will depend on the trend in the unfiltered series." As noted by Burns and Mitchell, filtering data is useful for studying secular trends, but has limited application in the study of classical cycles.

1.2.3 Using Quarterly GDP Data

In this research, GDP data for the U.S., U.K., and Australia are used to try and time cyclical turning points and model the aggregate economy. The data for the estimations and simulations are Gross Domestic Product in chained currency for the U.S. (1947Q1-2008Q4 (unrevised 2008Q4 estimate)), U.K. (1955Q1-2008Q4), and Australia (1959Q3-2008Q3). While there are similar measures of employment, production, and income for each of these countries, the main goal is to try and detect cyclical turning points and model any economy with data that are easily available. As has been pointed out by many researchers, we are trying to characterize and model *the* business cycle and not turning points in any individual time series. Using only aggregate GDP data to model output and detect turning points is the simplest approach, and the one that is followed here. The primary reason for using only GDP data is that this information is available for most countries on a quarterly basis, and therefore these methods can be applied broadly to characterize economic activity.

The use of quarterly data here also allows us to examine the usefulness of the Bry and Boschan quarterly algorithm developed by Harding and Pagan (2002) to describe cyclical features of different economies. Given that the algorithms developed by Harding and Pagan have proven to give reasonable estimates of turning points in the past provides us a straightforward method of measuring cyclical features that our models should be capable of matching in simulation.

1.2.4 The BBQ Dating Algorithm: Non-Parametric Turning Point Detection

Several studies have expanded upon the techniques of Burns and Mitchell (1946) developing methods that can be applied to any aggregate time-series data for the purpose of finding turning points. One of the primary uses of turning point detection methods is to help specify a model of aggregate activity that most accurately mimics the stylized facts observed in actual data. By using a consistent method to compare the features of simulated data from a number of models to underlying series, the work of Hess and Iwata (1997) and Harding and Pagan (2002) was able to eliminate a number of proposed models of the business cycle.

The work of Harding and Pagan (2002) adapted earlier work by Bry and Boschan (1971) to work with quarterly and unfiltered data ("the BBQ algorithm"). The initial algorithms written by Bry and Boschan were intended to detect turning points in monthly data series. However, the adaptation of this algorithm to quarterly data has proven to be a very good approximation to the NBER turning point dates in a quarterly sense. Additionally, Harding and Pagan do not filter the data used in their research for the reason that the trend and cyclical components cannot be independently determined or clearly separated.

Classical cycles are defined as those irregular periodic shifts between expansion and contraction in a time series. The classical cycle studied here is identified using the non-parametric BBQ method of determining the stylized facts or reference cycle displayed by an economy. Harding and Pagan (2002) show that the classical business cycle is a valid way to measure whether or not an economic model properly fits the data in question. Harding and Pagan (2002) use Monte Carlo methods to simulate time paths for several different business cycle models. Harding and Pagan find that many models designed to explain certain features of the business cycle actually are unable to recreate most features seen in the underlying data.

Econometric models are typically selected based on parametric measures of fit and tests for model specification using standardized tests minimizing residuals or forecasting error. Standard econometric measures of fit have no built-in method to examine the paths or features created by a particular model. However, in business cycle research if one ignores the fundamental features of the underlying series with regard to the length, depth, and shape of recessions or expansions a model may still be found to fit the data well statistically. In fact, the model may not be able to simulate time series that remotely resemble the actual cyclical features of the underlying data.

Using the methods of Harding and Pagan (2002) the reference cycles for the three countries are assembled using the "standard BBQ algorithm." The economic time series are assumed to exhibit phases that last at least k quarters, and full cycles that last at least m quarters. In what we term here as the "standard algorithm" the values of k and m are set such that $k = 2$ for both recessions and expansions and $m = 5$.

The requirement that $m = 5$ serves as a censor limiting detected peaks and troughs to be a certain minimum distance apart. The values of minimum phase length and full cycle requirements are intended to match the description of a recession by Burns and Mitchell (1946) that recessions should last six months at full cycles should last at least 15 months. Also, the turning points are censored such that peaks and troughs must alternate so that continuous cycles can be measured. Lastly, rather than use the typical rule of thumb "two consecutive quarters of negative growth signals a recession" a peak is found when $\Delta_2 y_t > 0, \Delta y_t > 0, \Delta y_{t+1} < 0, \Delta_2 y_{t+2} < 0$. In this specification the number of periods looking forwards and back is defined as the "window width" of $k = 2$. A trough is found when $\Delta_2 y_t < 0, \Delta y_t < 0, \Delta y_{t+1} > 0, \Delta_2 y_{t+2} > 0$ where the value for $\Delta_2 y_{t-2} = y_t - y_{t-2}$.

While the standard BBQ algorithm is useful at timing business cycles as it is currently specified, it would not capture a turning point where there is a single quarter of negative growth at $t+1$ if the economy returns to a higher level of output in $t+2$ relative to time t. A series would need to experience two consecutive quarters of negative growth or not return to a higher level of output at time $t + 2$ relative to output at time t in order to be characterized as a recession. This definition of a turning point assumes that the two quarters prior to the peak had levels of output lower than at the peak. In Chapter Three, the restrictions on the phase length k and the complete cycle length censor m are relaxed in the algorithm, and each simulation is retested to see if model performance is an artifact of the algorithm used to define cycles.

Growth that is below trend would not appear as a recession in output, and would fall under the umbrella of growth cycles. Our characterization of turning points in output is dubbed the *classical cycle* versus the *growth cycle* since it detects peaks and troughs in the level of output and not growth. As was previously mentioned, censoring mechanisms are employed in the algorithms in order to assure that complete cycles are being measured. The complete censoring mechanism used here and by Harding and Pagan (2002) is described by Cashin and Ouliaris (2004) and employed as follows…

1. Enforce alternating peaks and troughs.
2. Enforce peak-to-peak and trough-to-trough restriction on complete cycle length (i.e., $m = 5$).
3. Eliminate dates at the beginning and final two quarters of the series.
4. Eliminate phases with a duration shorter than the restriction (i.e., $k = 2$).
5. Again ensure that peaks and troughs alternate.

After peaks and troughs have been determined, the length of contractions (P→T "Peak to Trough") and the length of expansions (T→P "Trough to Peak") can be calculated. The amplitude between a peak and trough is calculated as the simple difference between the two levels of output in terms of 100 times the log difference. Cumulated amounts of output are calculated over the entire length of a recession or expansion as a percent change from the previous peak or trough. For example, if an economy has 5% growth during period $t+1$ and 0% growth in $t+2$ the flow of output at time $t+2$ is still 5% higher relative to time t yeilding cumulative gaines of 10% relative to time t. Lastly, percentage excess is calculated by measuring the difference between the *triangle approximation* of cumulated output and the actual path of the economy. A negative (positive) excess in a contraction implies that the actual data series decreased faster (slower) when compared to a straight line as displayed in Figure 1.1. It is possible that an excess measure of zero occurs if for example the economy grows very quickly from the trough of a recession, and then experiences slower than trend growth for a time, before experiencing much faster growth. While the excess measure is imperfect it does yield some insight on the asymmetry experienced in a business cycle.

Due to the fact that the BBQ algorithm is not perfect at matching NBER turning point dates, several authors have tried to fix this issue by calibrating the algorithm to be more exact. The BBQ algorithm is modified by Morley and Piger (2005) ("M&P" hereafter) who replace the zeros with an optimized value $\alpha \neq 0$ in order to remove the BBQ bias to the NBER turning point dates. However, it is unclear what bias you would be removing or how to calibrate the algorithm when no NBER dates exist where one extends the use of the modified algorithm for the cases of the U.K. and Australia GDP or any other series.

The modification to the BBQ method proposed by M&P improves the timing of turning point selection by pinpointing 14 of 19 NBER announced turning points versus 9 of 19 that are pinpointed using the standard BBQ algorithm. The turning points selected by the standard BBQ process are mostly only one quarter off from the NBER announced dates. Research by Camacho and Perez-Quiros (2007) showed that missing the NBER turning points by a single period might be problematic if one is trying to study the different behavior of an economy in a recession or expansion. The importance of removing the bias in the standard BBQ algorithm is uncertain, and comes at the cost of having to estimate parameters of an optimal algorithm based on an individual series. Further study of this possibly important modification is beyond the scope of this paper.

The primary use of the standard BBQ routine is for defining a chronology of events rather than assigning a conditional probability to being in a recession at a given moment in time. In comparison, even the methods of Chauvet and Hamilton (2005), and Chauvet and Piger (2003) which use real-time rather than revised data, can only assign probabilities to business cycle turning points several months after the fact.

Many models have been developed over the past twenty years that have attempted to recreate, simplify, clarify, or replace the NBER dating methods. Using both univariate and multivariate frameworks with both linear and non-linear structures. Turning point dating methods used by McConnell and Perez-Quiros (2000) which built upon Hamilton (1989) provide a very close approximation to the NBER dates, but use quarterly data to recreate the turning points so are unable to pinpoint the NBER turning points which are officially declared in months. Recent work by Leamer (2008) offers potential improvement for multivariate dating algorithms, as he uses monthly data on unemployment, employment, and production to define a recession-dating algorithm. Leamer finds that his algorithm nearly completely approximates the NBER turning points, missing only a single date. However, Leamer states in May 2008 that things would have to get much worse before declaring a recession. Unfortunately, Leamer's algorithm mis-times the December 2007 recession displaying the difficulty of forming an algorithm that can perfectly mimic the NBER's methods, which the NBER even admit are not fixed rules (National Bureau of Economic Research, 2001).

The advantage of the non-parametric BBQ method presented here is that we can easily determine if and when we have experienced a turning point in hindsight using a simple algorithm. While the true method of business cycle dating obviously involves an amount of judgment on the behalf of the NBER-BCDC, it seems clear in the literature that non-parametric methods of business cycle dating have significant value if they are applied universally. Methods of determining the best macroeconomic models have been put through a battery of tests attempting to help with the model selection process. The actual economy, and the reference cycles that are displayed, should be sufficiently mimicked in data simulated from a model whose intent is to understand cycles. One of the goals of this research is to merge the use of non-parametric dating methods with the parametric dating and modeling methods that should agree on many levels. Until the NBER turning point dating method changes to a fixed algorithm, it might be impossible to match their logic for choosing certain turning points.

1.2.5 Business Cycle Characteristics

Fitting the features of the classical business cycle is a challenge for any model, linear and non-linear alike. The results of this paper show that the classical business cycle is almost undetectable in the U.S., U.K., and Australia during the Great Moderation using the standard BBQ algorithms. In order to properly compare the macroeconomic models suggested by various authors, a suitable set of stylized facts regarding the actual economy must first be explored. After the stylized facts are presented, the estimation models are presented, and then simulated to test their performance at simulating the stylized facts.

Table 1.2 displays the result of using the standard BBQ algorithm ($k = 2, m = 5$) to detect turning points and cyclical characteristics for the U.S., U.K., and Australia. In the U.S., the BBQ algorithm with a phase length of two quarters and a total cycle length of five quarters, ten peaks are detected. The number of peaks differs from the NBER's eleven peaks, as the 2001 recession is missed by the BBQ algorithm. According to the algorithm, a typical U.S. recession is 2.9 quarters long, while the typical expansion is 21.7 quarters. Compared to the NBER's 3.3 quarter long recessions and 19 quarter expansions, the algorithm appears very close on most levels. As the 2001 recession is not included in the BBQ summary, some of the difference between these numbers can be easily explained. Figure 1.2 shows the different periods where recessions are detected by both the NBER and BBQ algorithms. In Figure 1.2 it is sometimes difficult to discern between BBQ recessions and NBER recessions because they are often concurrent.

Figure 1.3 examines the 2001 recessionary period in closer detail to show why the algorithm fails to detect this mild recession by most standards. The 2001Q2 period would have been declared a business cycle peak, and 2001Q4 a trough, had the 2001Q4 GDP measure not exceeded the GDP measure in 2001Q2. As this point is not declared a peak according to the algorithm, the ensuing recessionary period and recovery is included in the same expansion that occurred prior to 2001. This very long recovery serves as the main source of increasing the average expansion length reported by the BBQ algorithm above the official NBER measure. Only once the 2008Q2 recession is detected in U.S. GDP data can the complete cycle be measured. Future revisions to output data may change the date of the most recent turning point declared by the algorithm.

Figure 1.4 shows that the BBQ algorithm will occasionally miss an official NBER turning point by a period or two. In the case of the 2008Q2 recession, the algorithm does not detect a peak in business cycle activity until two quarters after the NBER officially declared the 2007Q4 peak which signaled the beginning of the most recent recession. The error is made here because of a situation similar to what is seen in Figure 1.3 as GDP rose to a higher level in 2008Q1 and 2008Q2 relative to 2007Q4 even though the economy continued to be in distress. Outside of the peak which was completely missed, six of ten NBER peak turning points examined here were identified correctly. The peak in 1957Q3 was one found one quarter after the NBER date, and the peaks in 1960Q1 and 1969Q3 were located one quarter before the NBER date. Regardless of the inability to synchronize the turning point dates of the standard BBQ algorithm with the NBER turning point dates, there is little universal guidance for countries other than the United States.

1.2.6 Parametric Business Cycle Dating

An alternative to using a non-parametric algorithm to date business cycles is to use a parametric model that can estimate probabilities of being in a recessionary state during a given period. Hamilton (2003) criticized the work of Harding and Pagan (2003b) for using business cycle dating algorithms that mimic the NBER dating committee decisions to influence their model selection process. Instead of using an algorithm to detect turning points, Hamilton suggests using a Markov-switching model which is able to use observed data to create predicted states of the economy. Chauvet and Piger (2003) and Chauvet and Hamilton (2005) specifically show that MS models can accurately predict NBER business cycle turning points for U.S. data. However, this line of research depends critically on the ability to match the NBER turning point dates to determine the specification of the model. While it seems that it would be ideal to use some parametric model to estimate the probability of a recession, the choice of the model is also constrained to the current set of knowledge regarding the system. McConnell and Perez-Quiros (2000) show the basic Hamilton Markov-switching model in mean does not detect mild recessions like those since the Great Moderation unless the model is modified to switch in variance as well.

The Markov-switching approach assigns probabilities to the state of the system (i.e., recession/expansion) and has the advantage of being parametrically determined from the data even though the state is not directly observed as shown by Hamilton (2005) and Kim, Morley, and Piger (2005). However, MS models are vague on determining the actual timing of a turning point with the general rule being that the economy is in a recession when the probability of being in a recession is greater than 0.5. In many of the estimations performed here and elsewhere, the estimated rate of growth during a recession is positive. Another drawback of the MS framework is the model selection process which often simply refers to a linear model to determine lag length, dummy variable presence, or volatility switching.

Hamilton (1989) was the first to use the Markov-switching approach to determine business cycle turning points. The original MS model estimated by Hamilton was designed to estimate when the economy was in a high-mean growth state (expansion) and a low-mean growth state (contraction) using a non-linear filter. The MS approach to modeling business cycle turning points has since been expanded upon a great deal, including variance switching and VAR versions. Krolzig and Toro (2005) specifically use a VAR approach to study European business cycle features similar to those examined here. McConnell and Perez-Quiros (2000) expanded the MS model with a variance term that was able to switch independently of the mean growth rate. McConnell and Perez-Quiros (2000) showed that the MS model could predict business cycle turning points even in a period of low-volatility. As in previous studies a recessionary state is predicted from MS models when the probability of being in a recessionary state is greater than 50%. This implies that predicted "recessionary" states are only statistically likely and not certainties.

As is shown in Chapters Two and Three, the MS approach can often give very good predictions of turning points using various specifications, and yet give very poor predictions with other specifications that are statistically preferred to those with simulations that more closely resemble the data. In our paper, we merge these two lines of research together to help select a Markov-switching model that is able to match the business cycle statistics gathered using non-parametric methods.[1]

1.3 The Great Moderation and Modeling Cycles

The Great Moderation was first identified by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) as the simultaneous reduction in volatility of aggregate output and other macroeconomic variables in the mid 1980's. The Great Moderation has been studied in depth with most research debating whether or not the cause was good policy, good luck, a widespread technology change, or something entirely different. While the ultimate cause of the downward change in volatility has not yet been decisively found, there is little debate that some sort of moderation did actually occur. Whether or not a return to higher volatility can be prevented, or if this is even desired remains to be seen. Significant new evidence is presented in Chapter Two that shows that the Great Moderation may in fact be ending. The possibility that the Great Moderation is ending should help in identifying the possible causes for the reduction in volatility in the first place.

McConnell and Perez-Quiros (2000) and Kim and Nelson (1999) provide evidence that a Great Moderation of output volatility occurred in the United States around 1984 as is visible in Figure 1.5. McConnell and Perez-Quiros (2000) ("MPQ" hereafter) tested a simple linear model of GDP growth for a break in volatility and showed substantial evidence of a structural change. The Great Moderation occurred in many

[1] The merits of parametric versus non-parametric dating methods are discussed in papers by Harding and Pagan (2002); Hamilton (1989); Harding and Pagan (2003a); Hamilton (2003); Harding and Pagan (2003b); Chauvet and Hamilton (2005); Hess and Iwata (1997); Leamer (2008).

countries across the world at approximately the same time. As is visible in Figure 1.6 and 1.7, the Great Moderation occurred in the U.K. and Australia sometime in the early 1990s and mid-1980s respectively.

Our research examines the impact of the Great Moderation in the context of the measurement of classical cycles of output for the U.S., U.K, and Australia. The existence of the Great Moderation's reduced volatility has had a significant impact on the ability of traditionally measured classical cycle dating methods to detect peaks and troughs in the business (or reference) cycle. The issues laid forth in Section 1.2.5 show that during the Great Moderation the standard BBQ algorithm has performed poorly in detecting turning points as seen in Figures 1.4 and 1.3. Ultimately, it is found that linear models that account for the Great Moderation create significantly different cyclical features than those models that do not account for a downward break in output volatility. Additionally, statistical tests are used to determine the fit of simulated data from a variety of models when compared to the baseline cyclical features. Similar tests of simulated Monte Carlo data has been performed by Cogley and Nason (1995), Simkins (1994), and Hess and Iwata (1997) to help determine if a model fails at fitting the reference cycle's stylized facts presented by a dating algorithm.

1.3.1 Combining Parametric and Non-Parametric Methods

The model selection process for business cycles, has traditionally followed at least two distinct paths. The parametric path of model selection that many researchers have followed selects a model based on how well the parameters of a model fit the data. When using the traditional parametric fit measures to select a model, a researcher might test various specifications based on in-sample fit statistics to choose which parameters most closely fit the data. Tests such as a Likelihood Ratio test can help determine which model would be more likely to represent an underlying data series in a statistical sense. Parametric selection for a business cycle model might also use the fit of an out-of-sample forecast with the purpose of selecting a model that can provide the best forecasts of the future state of the economy.

The methods of mechanically defining turning points in economic time series began with Burns and Mitchell (1946) and was later adapted by Bry and Boschan (1971) who first dated turning points using a computer algorithm. Later work by King and Plosser (1994) examined Real Business Cycle ("RBC") models within the context of the test of the Adelmans using the turning point algorithms developed by Bry and Boschan (1971). King and Plosser (1994) concluded that since the RBC models they tested and the Keynesian model tested by Adelman and Adelman (1959) led to similar conclusions that the dating algorithms of Burns and Mitchell (1946) might be an artifact of their dating algorithm.

Rather than simply looking at parametric statistics regarding fit and forecasting ability, recent articles by Harding and Pagan (2002) Harding and Pagan (2003a) and Hess and Iwata (1997) have shown a revived interest in measuring the quality of business cycle models by using Monte Carlo methods to measure which specifications fit the data better with regard to the features of business cycles. As shown by Harding and Pagan (2002) and Hess and Iwata (1997) it is possible to test linear versus non-linear modeling methods in their ability to simulate data that resembles what is observed by combining parametric and non-parametric methods of model selection.

Rather than rely on selecting a specification based solely on parametric measures of fit or minimum forecasting error Harding and Pagan (2002) show that the classical business cycle is a valid way to measure whether or not an economic model properly fits the data in question. Non-parametric measures show that if one ignores the fundamental features of the underlying series with regard to the length, depth, and shape of recessions (expansions) a model is in danger of being assumed to fit the data rather well when it may in fact not be able to reproduce time series that remotely resemble the actual levels or features of output.

The issue at hand here is whether or not the classical cycle dating methods employed by Harding and Pagan (2002) or Hess and Iwata (1997) are useful in the context of the Great Moderation. McConnell and Perez-Quiros (2000) use a relatively simple macroeconomic model for output (ARIMA(1,1,0)) to show that a break in volatility occurred around the beginning of 1984. With the justification of a variance break, the authors adapted a Hamilton (1989) Markov-switching model in mean only to account for the downward break in volatility. More recently, there has been widespread study of the Great Moderation using VAR models with Markov-switching mean and volatility with very mixed results.

The results of Chapter Two and Three show much promise for the MS approach to modeling business cycles, as these models are able to mimic observed cyclical statistics in simulation. These results hold true even when the parameters of the algorithm are allowed to change to reflect the shorter and smaller recessions experienced during the period of the Great Moderation. These findings also show that we should compare richer linear and non-linear models to a linear model which accounts for a structural change in volatility like that put forth by McConnell and Perez-Quiros (2000).

1.3.2 Testing Simulations Against Stylized Facts

In order to test model specifications against one another, two-phase process is employed here. In the first phase we estimate the parameters for each model and collect measures of statistical fit for each specification. The followers of the parametric fit approach tend to stop here, and select the model that has the highest, most statistically significant measure of fit. Other researchers take the two-phase approach and use the estimated parameters to create Monte Carlo paths. Typically researchers make the assumptions of no parameter instability and normally distributed error terms. These assumptions which are used in this paper, serve to simplify the Monte Carlo process, but could be relaxed if desired. In future work it would be useful to test the importance of relaxing these assumptions on the model selection process by either using jackknife or bootstrap techniques to create parameter instability and non-normally distributed residuals for re-sampling.

Cogley and Nason (1995) examined Real Business Cycle ("RBC") models in a similar manner to King and Plosser (1994), but instead tested many models for their ability to create simulations with autocorrelations and impulse response functions similar to those seen in the data. Cogley and Nason (1995) added statistical testing of the autocorrelation and impulse response functions by applying generalized Q-tests to the simulated data. Simkins (1994) tested RBC models using statistical methods similar to Cogley and Nason to measure the similarity of simulated data to the classical business cycle patterns first studied by Burns and Mitchell. While both King and Plosser (1994) and Simkins (1994) findings offered mixed support for RBC models, Cogley and Nason (1995) were able to show that RBC models lacked an internal propagation mechanism. The lack of internal propagation meant that RBC models relied on shocks to replicate cycles, leading to one of the harshest criticisms of the theory. Other work by Neftci (1993) measured the fit of simulated models with a "shape" to determine if the simulations were sufficiently close to the same specific cycle. Neftci (1993) examines data similar to that studied by Burns and

Mitchell (1946) finding that simple random walk models are more likely to simulate the shapes present in the actual economy relative to a richer ARMA specification.

Morley and Piger (2005) and Galvão (2002) specifically tested linear versus non-linear models, but neither offered statistical tests other than 80% empirical distribution functions for individual features of the data. The testing strategy employed here is similar to that employed by Hess and Iwata (1997) and Cogley and Nason (1995). Multiple business cycle features are tested simultaneously using generalized Q-tests providing evidence on the ability of models to produce statistically similar features relative to those seen in the actual data. Additionally, Kolmogorov-Smirnov ("KS") tests are used here to show that the individual features of ARIMA models including volatility shifts are significantly different than models that do not account for the Great Moderation. All features are measured in both the simulated and actual data using the standard BBQ algorithm in Chapter Two to identify turning points. In Chapter Three we relax the assumptions of the algorithm to analyze the robustness of its parameters.

1.4 Comparison to Previous Research

In our research we use a univariate framework to explore simulated data from a variety of linear, structural change, and non-linear MS models. By comparison, previous work shows that univariate MS models are generally superior to almost all other univariate business cycle models as measured by the fit of simulated data to actual cyclical features. It should be noted that the ARIMA(1,1,0) approach has been found to be comparable to a simple MS approach at replicating most cyclical features found in U.S. data, with the exception of asymmetry. Asymmetry is the condition that recessions are generally linear, while recoveries tend to show a rapid growth at first with slower growth once the expansion is mature. The inability for simple models to replicate complex features like asymmetry is noted in the work of Morley and Piger (2005), Clements and Krolzig (2004), Kim, Morley, and Piger (2005), Harding and Pagan (2002), and Hess and Iwata (1997). Morley and Piger (2005) note that they are not testing for model specification in their MS models, but generally testing for the importance of non-linearities. Nearly all previous literature on measuring the cyclical features of simulated MS data includes a comparison to an ARIMA model. These results provide statistical evidence that researchers should account for the Great Moderation when making comparisons between MS and ARIMA models. The approach taken here is to compare the simulations made by a range of models in the presence of the Great Moderation. One of the most important conclusions of our work here is that any MS models should be compared to a structural change in variance ("ARIMA-SV") model, since a simple linear model is incapable of replicating most business cycle features for any of the three economies examined.

The Great Moderation has had an important impact on the field of business cycle modeling. Foremost, the discovery of the Great Moderation did not occur until 1999, at which point most research used a constant measure of volatility in their modeling and simulation methods. Furthermore, estimates by Clements and Krolzig (2004) and Hess and Iwata (1997) are from a more limited 1949-1992 time frame prior to study of the Great Moderation. Harding and Pagan (2002) analyzed data through 1997 but failed to account for a break in volatility in their modeling approach. Kim, Morley, and Piger (2005) accounted for the Great Moderation when testing MS models over the 1949-2003 period, finding that volatility shifts were important to correctly fitting the model. However, Kim, Morley, and Piger (2005) did not include tests on simulated data for expansionary business cycle features or compare their results to a linear model which could account for the Great Moderation like the ARIMA-SV.

Results here show that simple ARIMA-SV and MS models that account for an independent shift in volatility are in fact very successful at replicating business cycle features in U.S., U.K., and Australian GDP. In fact in many cases an ARIMA-SV produces simulations that are comparable to those produced by a MS model. Thus, it is found that other researchers studying the dynamics of simulated data from business cycle models should be comparing their estimates to a model accounting for the Great Moderation like the ARIMA-SV specification. Also, if using a MS model to mimic the business cycle, one should take the Great Moderation into account and specify their models appropriately.

Table 1.1: NBER Business Cycle Expansions and Contractions

Business Cycle Reference Dates		Duration in Months (Quarters)							
		Contraction		Expansion		Cycle			
Peak	Trough	Peak to Trough		Trough to Peak		Trough from Previous Trough		Peak from Previous Peak	
Quarterly dates are in parentheses									
November 1948(IV)	October 1949 (IV)	11	(3.7)	37	(12.3)	48	(16.0)	45	(15.0)
July 1953(II)	May 1954 (II)	10	(3.3)	45	(15.0)	55	(18.3)	56	(18.7)
August 1957(III)	April 1958 (II)	8	(2.7)	39	(13.0)	47	(15.7)	49	(16.3)
April 1960(II)	February 1961 (I)	10	(3.3)	24	(8.0)	34	(11.3)	32	(10.7)
December 1969(IV)	November 1970 (IV)	11	(3.7)	106	(35.3)	117	(39.0)	116	(38.7)
November 1973(IV)	March 1975 (I)	16	(5.3)	36	(12.0)	52	(17.3)	47	(15.7)
January 1980(I)	July 1980 (III)	6	(2.0)	58	(19.3)	64	(21.3)	74	(24.7)
July 1981(III)	November 1982 (IV)	16	(5.3)	12	(4.0)	28	(9.3)	18	(6.0)
July 1990(III)	March 1991(I)	8	(2.7)	92	(30.7)	100	(33.3)	108	(36.0)
March 2001(I)	November 2001 (IV)	8	(2.7)	120	(40.0)	128	(42.7)	128	(42.7)
December 2007 (IV)				73	(24.3)			81	(27.0)
Average, all cycles:									
1945-2001 (10 cycles)		10	(3.3)	57	(19.0)	67	(22.3)	67	(22.3)

Source: National Bureau of Economic Research (2009)

Note: All information contained in this table is compiled by the NBER-BCDC, and is reprinted here with permission of the NBER. Quarterly recession statistics in this table are approximated by the author, by dividing the monthly statistics by three. Official NBER quarterly peaks and troughs are determined independently of monthly turning points which is why in some rare cases the two do not appear to be consistent.

Table 1.2: Business Cycle Peaks and Characteristics

	U.S.		U.K.		Australia	
	Year	Quarter	Year	Quarter	Year	Quarter
Cycle Peak 1	1948	4	1955	3	1960	3
Cycle Peak 2	1953	2	1961	2	1965	3
Cycle Peak 3	1957	3	1973	2	1971	3
Cycle Peak 4	1960	1	1974	3	1975	2
Cycle Peak 5	1969	3	1979	2	1977	2
Cycle Peak 6	1973	4	1990	2	1981	3
Cycle Peak 7	1980	1	2008	2	1990	1
Cycle Peak 8	1981	3				
Cycle Peak 9	1990	3				
Cycle Peak 10	2008	2				

Cycle Characteristics

	U.S.		U.K.		Australia	
	P→T[a]	T→P[b]	P→T	T→P	P→T	T→P
Durations	2.8889	21.7000	4.1667	31.0000	3.5714	16.5000
Amplitudes(%)	-2.2380	22.1355	-2.6847	24.2505	-1.7911	20.0444
Cumulative(%)	-2.9379	402.8080	-7.1992	522.5736	-3.6672	210.2755
Excess(%)	0.0244	1.1160	-0.0608	-0.5617	0.0737	0.7977

All measures are in quarters

[a]P→T refers to recessionary characteristics.

[b]T→P refers to expansionary characteristics.

Figure 1.1: Defining the Classical Business Cycle

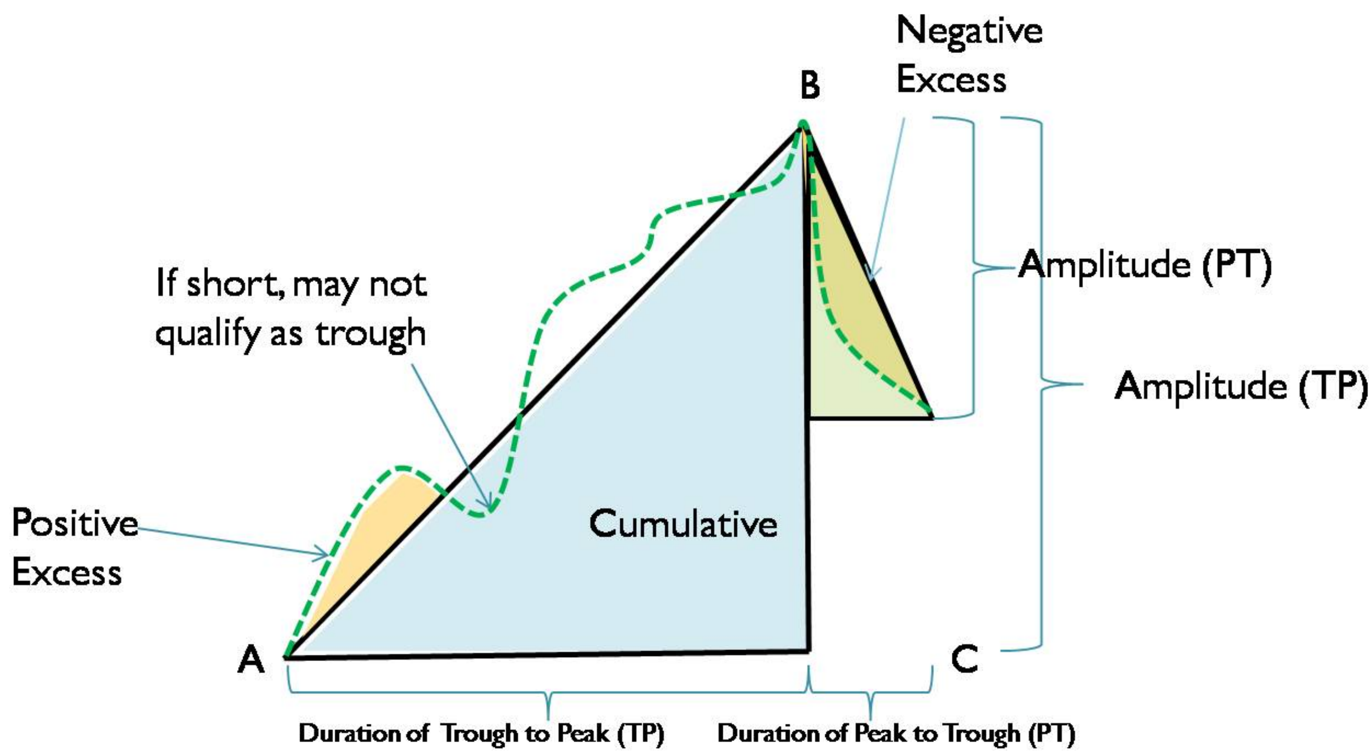


Figure 1.2: U.S. BBQ Algorithm Versus NBER Recession Dates

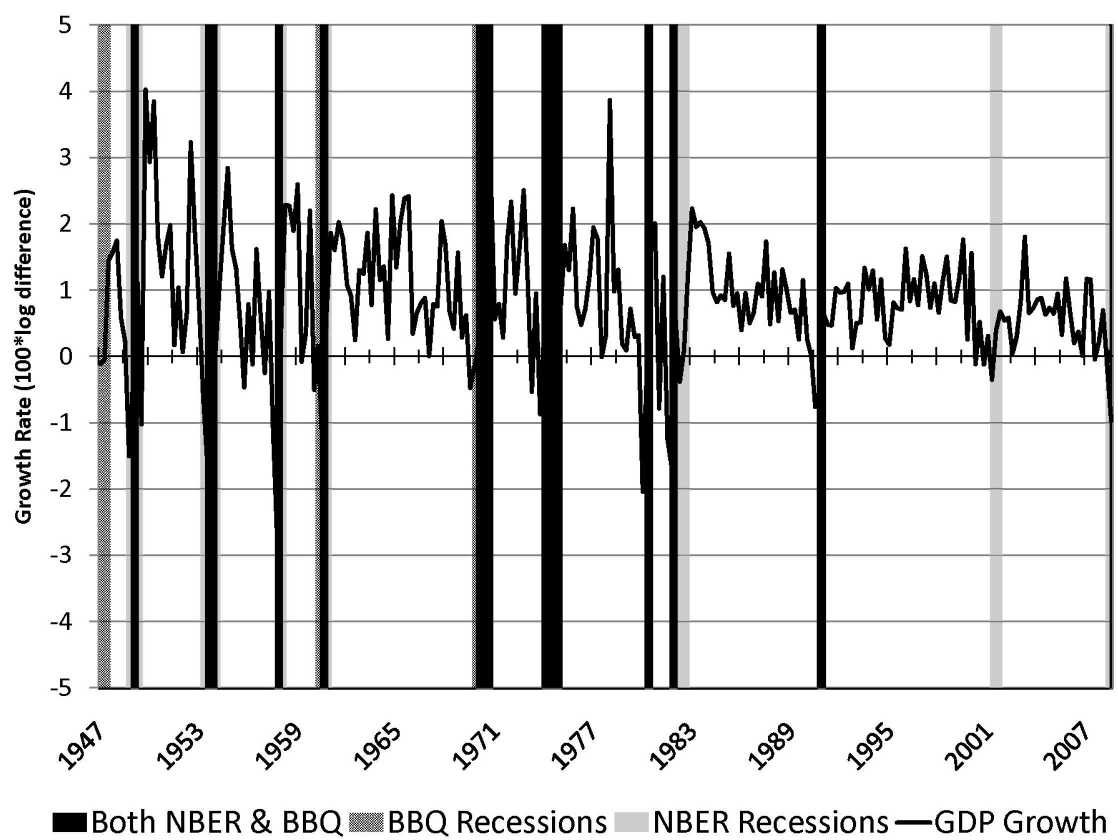


Figure 1.3: BBQ Algorithm Can Miss Turning Points

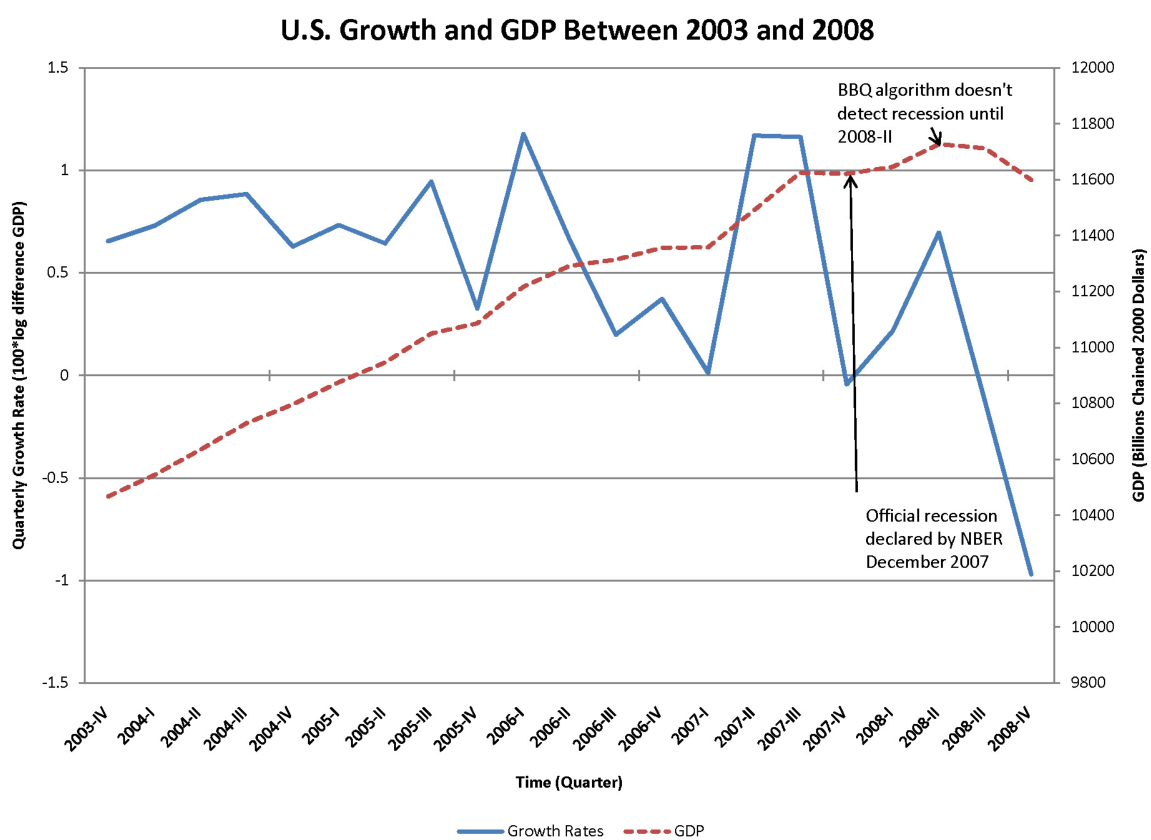


Figure 1.4: BBQ Algorithm Can Be Out of Sync

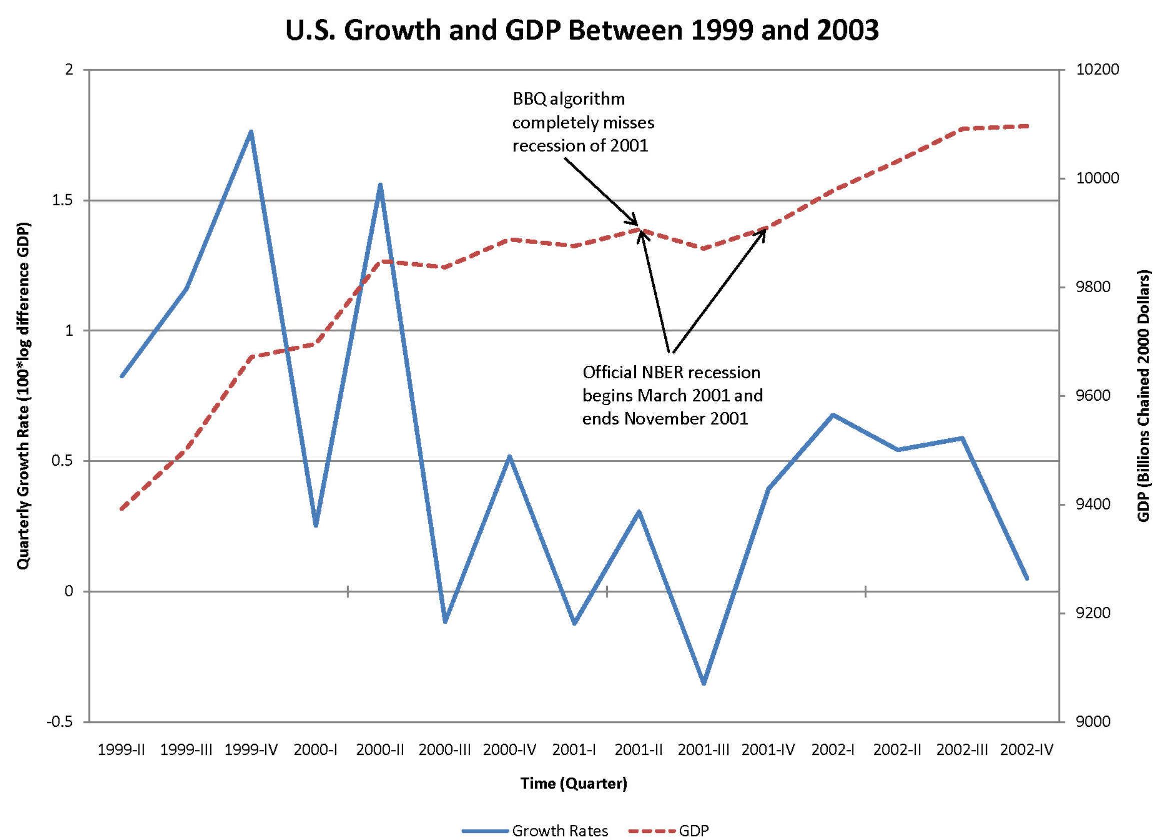


Figure 1.5: U.S. Growth Rates
(Gray Bars Indicate BBQ Identified Recessions)

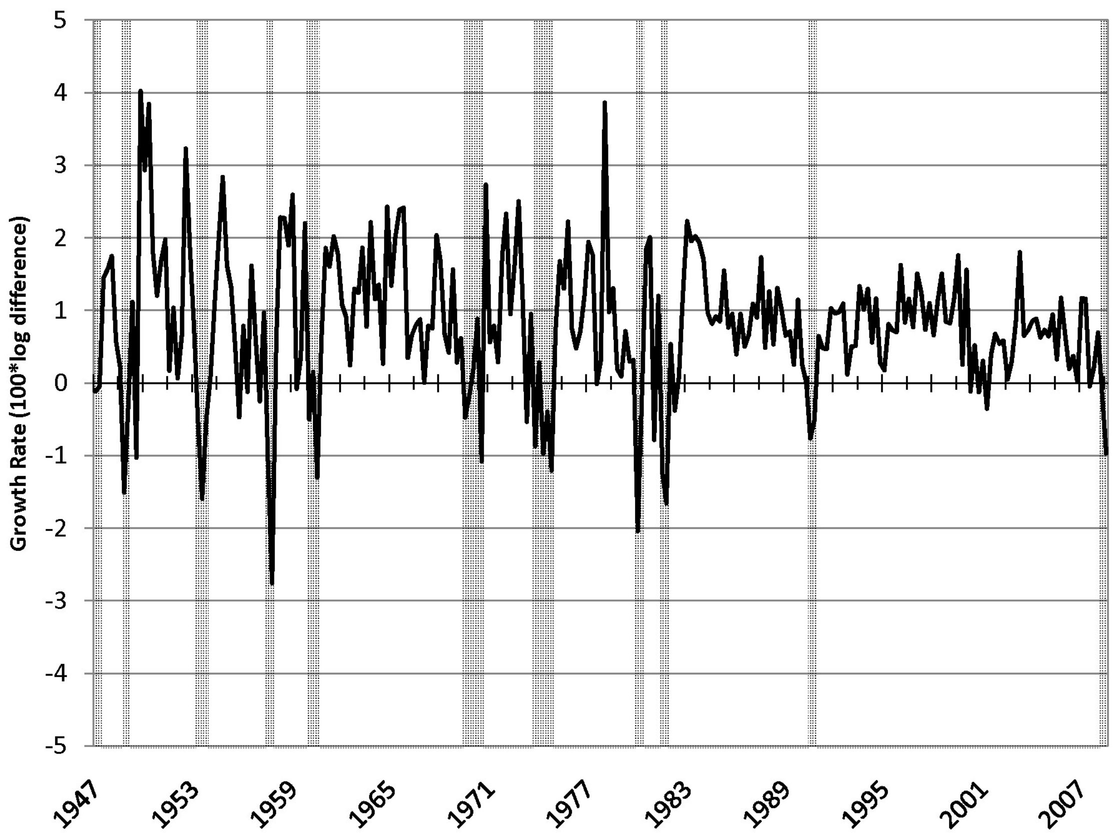


Figure 1.6: U.K. Growth Rates
(Gray Bars Indicate BBQ Identified Recessions)

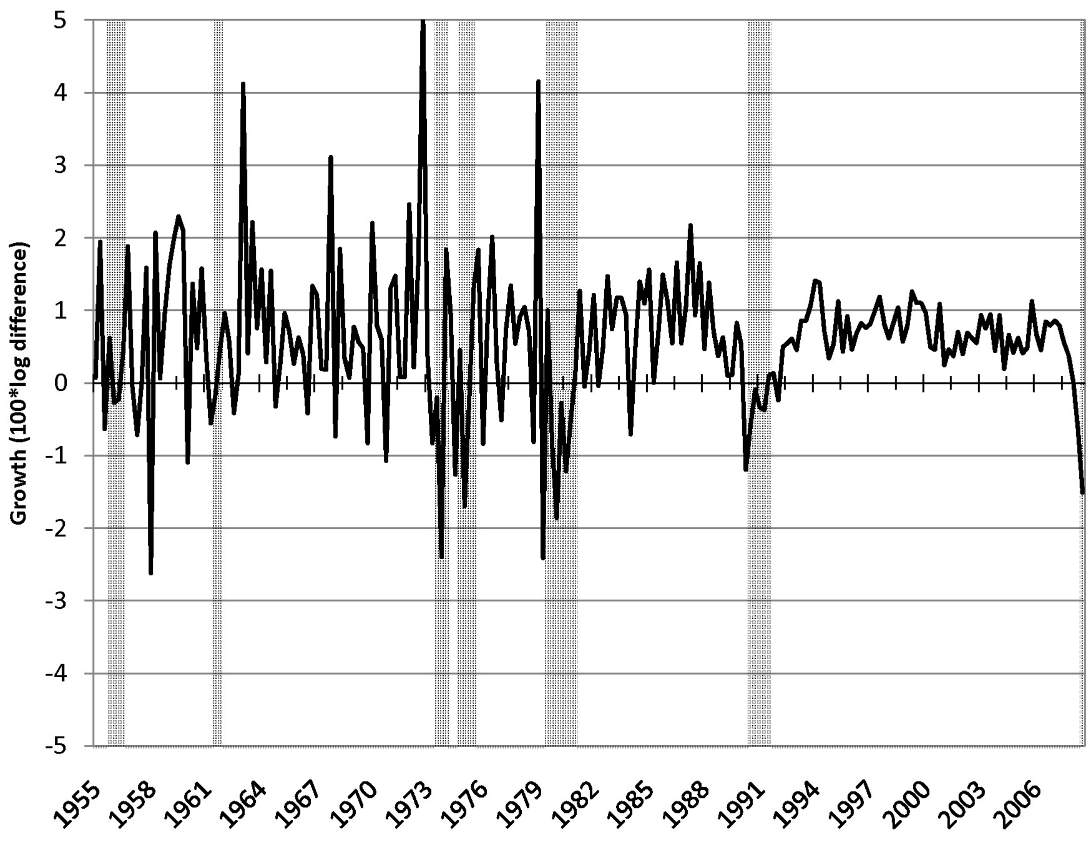
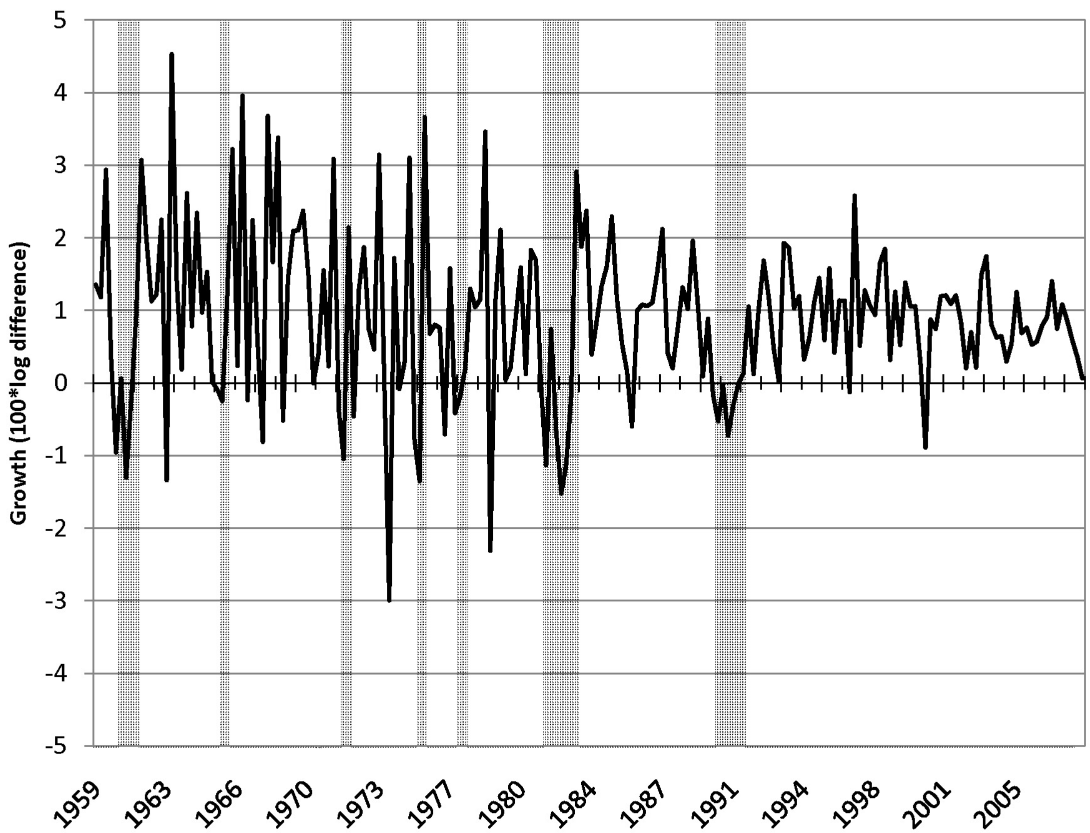


Figure 1.7: Australian Growth Rates
(Gray Bars Indicate BBQ Identified Recessions)



Chapter 2

Incorporating the Great Moderation Into Business Cycle Modeling

2.1 Introduction

The methods of determining peaks and troughs of business cycles are neither uniform nor transparent. Actual U.S. business cycle turning points are determined in hindsight by the National Bureau of Economic Research's Business Cycle Dating Committee ("NBER" hereafter). Macroeconomic models are frequently constructed to mimic NBER turning points in order to measure historical trends or form forecasts. Many models designed to mimic historical trends in output are deemed statistically sound. However, the same statistically sound models can often be dismissed due to their inability to replicate business cycle features in simulated data. Harding and Pagan (2002) and Hess and Iwata (1997) provide significant evidence that many valid statistical models perform very poorly in simulation experiments which test the ability of various specifications to replicate typical business cycle features. Recent work directly comparing simulated data to business cycle turning points has led authors such as Clements and Krolzig (2004) and Camacho and Perez-Quiros (2007) ("CPQ" hereafter) to discredit model specifications that might otherwise appear statistically valid. This paper tests a range of model specifications for their ability to simulate the business cycle characteristics measured in the underlying data series.

This research specifically examines the ability of linear, structural change, and non-linear Markov-switching models to simulate business cycle characteristics. Business cycle characteristics are measured with metrics commonly seen in the literature such as the length or depth of a recession. Using U.S., U.K., and Australian GDP data it is found that models accounting for the Great Moderation in output volatility significantly increase the ability to replicate business cycle features. This research also provides supporting evidence to CPQ that linear structures in MS models are not necessary for replicating business cycles. Growth in the U.S., U.K., and Australia generally behave according to a jump-and-rest pattern described by CPQ for the case of the United States. Our research is also in agreement with the work of Kim, Morley, and Piger (2005) and Morley and Piger (2005) ("M&P" hereafter) who found little need for linear structures in their non-linear models of the business cycle. As CPQ point out, a jump-and-rest business cycle implies that recessionary shocks do not have long lasting effects. This evidence is in contrast to much previous research that finds economic growth models exhibit autocorrelation. This finding is important to policy, as responses to economic shocks are often meant to limit shock persistence.

Markov-switching models initially proposed by Hamilton (1989) were able to describe an economy where growth rates jumped between "boom" and "bust" states.

Until the mid-1990s, MS models showed consistent ability to predict recessionary states using only the single macroeconomic GDP series. Typically, these MS models included autoregressive parameters intended to capture the persistence of shocks to the model. Following the Great Moderation however, MS models with constant variance ("CV") were unable to accurately predict recessions. The basic MS model was modified by McConnell and Perez-Quiros (2000) ("MPQ" hereafter) to incorporate the Great Moderation by showing that when the volatility state was were allowed to shift independently of the growth state, the MS model was again able to predict recessions. Since the publication of MPQ however, many studies employing the MS model have used a simpler specification where volatility states are tied to the growth state. Thus, when the model is in an expansionary period, it would have low-volatility and vice-versa. This paper shows that the simpler CV or tied switching ("TS") volatility models are not a substitute for independent switching ("IS") volatility.

The results presented in this chapter are significant as many other researchers continue to use MS models that are improperly specified, believing that they are accounting for the Great Moderation. Additionally, many researchers justify the use of Markov-switching by comparing their models to a linear model without a structural change in volatility. Any results where the MS model is compared to a simple linear structure are very likely to be misleading, as it is shown that these simple models are unable to mimic most business cycle characteristics with any regularity. Structural-change in volatility models of GDP growth are very successful at simulating most business cycle features, and often outperform many models containing Markov-switching. This finding is important as structural-change models are intuitively appealing due to the fact that they are much more straightforward than MS methods.

In this paper, evidence is provided supporting the idea that business cycle dating algorithms are useful in helping select an appropriately specified linear or MS model for any series when trying to simulate cyclical features of data. Using a combination of the MS approach and simple business cycle dating algorithms results in selecting models that are superior for the purpose of explaining certain features of the data. The Great Moderation of output in the U.S. and other industrialized economies has spawned a great deal of research in an attempt to identify the cause. Whether the impetus of the Great Moderation was good luck, good policy, or something different is irrelevant to the modeling strategy employed here. The preferred specification for the countries studied here predicts a possible end to the Great Moderation in U.S. and U.K. GDP. If the Great Moderation has ended, then only the models including independent switching would be able to detect this change.

The models used here are useful at gaining a better understanding of the stylized facts of business cycles even though they do not contain explicit optimizing agents or theoretical restrictions. An improved understanding of the nature and predictions of simple models will help more complex and multivariate models forecast behavior in an economy. There is a valued interest in being able to improve predictions about future business cycle peaks and troughs and knowing how long business cycles are expected to last. Recent global economic turmoil may be signaling a large shock, a series of large shocks, or even an end to the Great Moderation. Thus it stands to reason that economists should have a better understanding of the predictions made by our models.

2.2 Background on the Great Moderation and MS Modeling

Previous work studying business cycle models and the Great Moderation have similarly concluded that non-linear features are essential.[1] However, evidence presented here displays that the specific type of non-linearity is important to the ability of a model to fit stylized business cycles. Beginning with a simple linear structure for output like the ARIMA(1,1,0) model in Equation 2.1, a large variety of models can be formed and tested.

$$\Delta y_t = \mu + \phi \Delta y_{t-1} + \epsilon_t, \epsilon \sim N(0, \sigma^2) \tag{2.1}$$

Previous work by Morley and Piger (2005) tests for linear effects in the U.S. using MS models, concluding that little is added with a simple non-linearity. However, the models estimated by M&P do not allow the mean and volatility to switch states independently. Independent switching mean and variance yield the best fitting models tested here, while tied-switching models conform to the predictions of M&P. Kim, Morley, and Piger (2005) ("KMP" hereafter) estimated independent switching models for each of the countries studied here, and found they all fit the MS model well. However, KMP focused only on testing their models for recessionary features and not all business cycle characteristics. As we will see, tests here show that models with independent switching mean and variance can account for all business cycle features versus those models where either the variance is constant, or the mean and variance shift simultaneously. In fact for the U.S., U.K., and Australia, there is substantial

[1] See Smith and Summers (2002), Clements and Krolzig (2004), Kim and Nelson (1999), and Morley and Piger (2005)

evidence that the MS-TS models are able to replicate recessionary features but not expansionary features. This is an important finding as much previous research has estimated MS models with either constant volatility or shifting variance which is tied to the state of the mean (See Clements and Krolzig (2004), M&P, Galvão (2002), and Smith and Summers (2005)). This paper does not test the "bounceback" MS model proposed by KMP which is specifically designed to improve the model's ability to capture recessionary features. The work of KMP also differs from our results here as they account for the Great Moderation using a dummy variable set to 1984Q1, rather than letting the Monte Carlo simulation volatility switch between states according to a latent state variable.

Morley and Piger find that non-linearity in mean growth rates are important at replicating business cycle features, while shifts in volatility are not. However, the conclusions of M&P are based on additional features not tested here such as average growth rates and variances during different phases. Also, as noted on page 13, Morley and Piger use a modified dating algorithm making direct comparison to other studies unclear. Also, M&P do not allow volatility to switch by itself in their simulations, but instead force a switch at the estimated date of the Great Moderation. These differences make their results difficult to compare to those presented here.

Galvão (2002) tests non-linear and linear U.S. models for their ability to fit asymmetric features of the business cycle. Asymmetry as far as we are concerned here is the condition that recessions are generally linear, while recoveries tend to show rapid growth at first with slower growth once the expansion is mature.[2] Galvão finds that non-linear models are not generally flexible enough to capture asymmetric fea-

[2]See Morley and Piger (2005), Clements and Krolzig (2004), Kim, Morley, and Piger (2005), Harding and Pagan (2002), and Hess and Iwata (1997).

tures. The ARIMA(1,1,0) approach has been found by Hess and Iwata (1997) and Harding and Pagan (2002) to be comparable to a simple MS approach at replicating most cyclical features found in U.S. data, with the exception of asymmetry. This paper agrees generally with M&P and Galvão (2002) that certain MS models are better at replicating business cycle features such as asymmetry. However, many MS models estimated here cannot replicate business cycle characteristics with any accuracy. Generally simple MS models with CV or TS volatility perform much worse than linear structural-change in volatility models that account for a break like the Great Moderation.

Some have questioned the use of business cycle dating algorithms that mimic the NBER business cycle dating committee decisions.[3] Instead of using a dating algorithm, Hamilton (2003) suggests using the predicted states from a simple MS estimation to determine business cycle turning points. Chauvet and Piger (2003) and Chauvet and Hamilton (2005) specifically show that MS models can come very close to predicting NBER business cycle turning points for U.S. data. However, the merging of the parametric methods of Hamilton and the non-parametric methods of Harding and Pagan (2002) is regularly performed. A prime example of merging these methods is in KMP who employ both parametric and non-parametric approaches to justify the use of their bounceback MS model. In application to the U.K. and Australia, MS models have mixed success fitting business cycle characteristics. The U.K. and Australia experienced their own Great Moderation in output volatility in the early 1990s and mid-1980s respectively as seen in Figures 1.6 - 1.7. While the Euro area has a business cycle dating committee similar to the NBER in the Centre for Economic Policy Research, there is no worldwide standard for dating turning points. When

[3] See Hamilton (2003) for specific criticisms and Harding and Pagan (2003b) for a rebuttal.

trying to fit the MS models estimated here, a common non-parametric algorithm is used to detect turning points in output. This research supports the idea that business cycle dating algorithms are useful in helping select an appropriately specified linear or MS model for any series that is meant to mimic cyclical features of data. By combining the MS approach, and business cycle dating algorithms these results show that the U.S., U.K., and Australian business cycles can be best approximated using Markov-switching models with independent switching. In the case of the U.S. and Australia linear models with structural change accounting for the Great Moderation are also able to replicate nearly all features of the business cycle.

2.3 Testing Strategy

The testing strategy employed here was first used by Adelman and Adelman (1959) to test a large Keynesian-type macro model. The "test of the Adelmans" showed that despite the lack of internal propagation mechanisms, random shocks could create cycles in simulated data that resembled those in the actual economy. First, a number of models that are commonly used in the business cycle literature are estimated over the time period of available data. Each of the original time paths that are used in the estimation phase have their features measured using a standard BBQ algorithm as described in Section 1.2.4. Next, using the parameters that are estimated in each of the models, Monte Carlo paths are generated for each model assuming no parameter instability and normally distributed errors. For each of the simulated time paths the features are measured using the same algorithm used to measure the original time series. Finally, statistical tests are performed to estimate the likelihood of seeing similar features in simulated time paths that are seen in the original data.

Work by Hess and Iwata (1997) tested simple linear models of the economy against those with more sophisticated linear and non-linear dynamics. Specifically Hess and Iwata (1997) tested models designed to capture asymmetrical features of the data concluding that an ARIMA(1,1,0) model was equal to or superior to other models tested, including a basic MS model. Hess and Iwata (1997) used generalized Q-tests and Kolmogorov-Smirnov tests to demonstrate that there was little to be gained in using more sophisticated linear and non-linear modeling. Hess and Iwata (1997) used a "running peak" turning point algorithm that differed from that of Burns and Mitchell (1946) and Bry and Boschan (1971) making the results somewhat difficult to compare to those from other studies. Hess and Iwata (1997) also did not test their simulated data for asymmetric features that Neftci (1993) and Kim, Morley, and Piger (2005) showed to be present in U.S. data. The conclusions of Hess and Iwata are based on the ability of a model to replicate the cyclical nature of the historical data. In our research here Q-tests are utilized to compare the features of simulated data the same features in actual business cycles for each country. Kolmogorov-Smirnov tests are also used here to compare the distributions of individual features produced by a variety of models. In combination, these tests allow several models to be ruled out of contention to fit business cycles.

Later work by Harding and Pagan (2002) adapted the dating algorithms of Bry and Boschan (1971) for quarterly data ("the BBQ algorithm"), and tested simple linear models against MS models popularized by Hamilton (1989).[4] Harding and Pagan (2002) results agreed with Hess and Iwata (1997) that non-linear models offered little gain over linear models. However, the work of Harding and Pagan (2002) lacked statistical testing like that in Hess and Iwata (1997) and Cogley and Nason (1995).

[4]See page 11 for a more complete description of the algorithm.

The parametric approach (i.e., Markov-switching) of determining a turning point was defended by Hamilton (2003) in a published debate where Harding and Pagan (2003a) supported the use of non-parametric measures of fitting models to the data being referenced.[5] The method of applying an algorithm to data to determine turning points is now common in the literature. Non-parametric characteristics of actual and simulated data can be compared to see if the model can in fact simulate the features of the underlying data. While work like that of Morley and Piger (2005) has shown that the methods of Harding and Pagan (2002) can be improved on by calibrating the algorithms to more closely match the NBER decisions. However, there is no natural NBER dating tool that can be used for international research. All business cycle characteristics compared here are measured using the unaltered BBQ algorithms developed by Harding and Pagan (2002).[6]

It is possible that an algorithm that can fully replicate the NBER methods would be very important to this field of research. Recently CPQ showed that the NBER business cycle turning point dates were "unique" versus other turning point dates like those determined using an algorithm. When CPQ included dummy variables (N_t) indicating NBER recessionary states in a simple ARIMA model, the autoregressive component of the model (ϕ) was no longer significant (See Equation 2.2). In a search to see if the NBER dates were truly special, CPQ show that even lagging or leading the NBER dates by a single quarter led to a significant ϕ value. These results are promising developments in the search for defining business cycle turning point algorithms or methods. An analysis similar to CPQ is performed here to test for the

[5]For more details on the debate see Hamilton (2003); Harding and Pagan (2003a,b).

[6]The search for the optimal business cycle dating algorithm is still open for further study, and is beyond the scope of this paper. Both M&P and CPQ show methods for improving business cycle dating for the U.S., but similar methods for foreign countries have yet to be developed.

importance of recessionary states. Recessionary states predicted from MS models and the BBQ algorithm are tested for each country.

$$\Delta y_t = \mu + \phi \Delta y_{t-1} + \beta N_t + \epsilon_t, \epsilon \sim N(0, \sigma^2) \quad (2.2)$$

In most previous studies of U.S. business cycles that estimate simple ARIMA models, the value of ϕ is found to be positive and significant. The significance of ϕ has led many researchers to believe in the importance of previous shocks to the current state of growth. The work of CPQ suggests that there is little persistence from a single shock. Using MS models accounting for independent volatility switching, CPQ show that autoregressive lags are not necessary for the United States. Camacho and Perez-Quiros (2006) describe their findings as a "jump-and-rest" pattern of output growth. Our findings show supporting evidence for the findings of CPQ for the U.S., but conflicting evidence for the U.K. and Australia.

2.4 Methodology and Data

This paper first develops linear, structural-change in volatility, and MS estimates of the movement in GDP. The data for the estimations and simulations are seasonally adjusted Gross Domestic Product in chained dollars for the U.S. (1947Q1-2008Q4), U.K. (1955Q1-2008Q4), and Australia (1959Q3-2008Q3).[7] The volatility of each of these countries varies a great deal across nations, and it is apparent from Figures 1.5 to 1.7 that the volatility of output growth rates has moderated quite abruptly. The variety of models that have been used to fit each of these series has displayed that the volatility breaks are likely to have occurred almost concurrently, however there are

[7] 2008Q4 data for the U.S. and U.K. are unrevised estimates.

a range of estimates on the exact date of each country's moderation. Work by van Dijk, Osborn, and Sensier (2002) estimates the different possible confidence intervals for the break dates in volatility showing that many countries experienced breaks in output volatility during the mid-1980's to early 1990's.

Using the BBQ algorithm leads to the estimated business cycle characteristics seen in the bottom of Table 1.2. The typical U.S. "recession" length is 2.89 quarters (P→T), and results in a loss of 2.93% of accumulated output (See Table 1.2, column 1). Recessions also appear to be very close to linear with only 0.024% of net excess. Average U.S. "expansions" are around 21.7 quarters in length and experience gains of 402.8% in accumulated output on average. The U.S. expansions are in excess of the triangle approximation, showing that growth from a trough tends to be nonlinear in nature. Kim, Morley, and Piger (2005) test the ability for a "bounceback" model to improve on the fit of a MS model with respect to expansion from a trough period. While the "bounceback" MS model is not estimated here, our results support the findings in KMP. The standard BBQ algorithm finds that the U.S. has an estimated 10 business cycle peaks relative to the 11 estimated by the NBER. The U.K. and Australia are found to each have 7 cycle peaks over their sample periods.

2.4.1 Endogenous Break Tests and Linear Methods

Following the work by McConnell and Perez-Quiros (2000), Generalized Method of Moments is used to simultaneously estimate Equations 2.3 and 2.4 along with Equations 2.5 and 2.6 listed below. These two pairs of equations help determine if and when a structural change has occurred in the parameters of these models. This estimation method allows for endogenous determination of the break dates for the variance and other parameters. The choice for the variance estimation equation can be attributed to the methodology of MPQ who use an unbiased estimator of the standard deviation $\sqrt{\frac{\pi}{2}}\,|\epsilon_t|$ as per Davidian and Carroll (1987). Sensier and van Dijk (2004) also used a similar specification to estimate break dates in the volatility of 214 U.S. macroeconomic time series.

$$\Delta y_t = \phi_0 + \phi_1 \Delta y_{t-1} + \epsilon_t \tag{2.3}$$

$$\sqrt{\frac{\pi}{2}}\,|\epsilon_t| = D_{1t}\alpha_1 + D_{2t}\alpha_2 + \mu_t \tag{2.4}$$

The instruments used in the system are a constant, lagged growth, and the dummy variables indicating the regime. For each model, a Newey-West weighting matrix with 12 lags was used to calculate standard errors. Estimation was carried out using iterated GMM with instruments used to construct the initial weighting matrix and an optimal weighting matrix to calculate final parameters. Each model was estimated over the range of possible break dates (T) between $T_1 = 0.15n$ and $T_2 = 0.85n$ which are used to set the values for D_{1t} and D_{2t} in Equation 2.4. Values of $T_1 = 0.15$ and $T_2 = 0.85$ are set to avoid the problems of detecting breaks at the beginning and ends of samples. The dummies for volatility regime are set such that

$$D_{1t} = 1 \text{ if } t < T$$

$$D_{1t} = 0 \text{ if } t \geq T$$

$$D_{2t} = 0 \text{ if } t < T$$

$$D_{2t} = 1 \text{ if } t \geq T$$

The model is then estimated for each possible T value between T_1 and T_2, and Wald or Likelihood Ratio (LR) tests are performed for each of the estimated models. Due to the presence of a nuisance parameter (t), that exists only under the alternative, the distribution of the test statistics are non-standard. The test statistics used to determine whether or not a series contains a break are of the *average exponential* variety. Specifically, the *sup*, *ave*, and *exp* tests are performed across all estimations as suggested by Andrews (1993) and Andrews and Ploberger (1994) using the p-value correction suggested by Hansen (1997).

It is also plausible that given the simple system of equations that have been laid out, the model could have experienced a break in the lag coefficients or the constant. In order to plausibly rule out this possibility, regressions were run on the following system as well.

$$\Delta y_t = D_{1t}\phi_0 + D_{2t}\phi_1 + D_{1t}\phi_2\Delta y_{t-1} + D_{2t}\phi_3\Delta y_{t-1} + \epsilon_t \tag{2.5}$$

$$\sqrt{\frac{\pi}{2}}\,|\epsilon_t| = D_{1t}\alpha_1 + D_{2t}\alpha_2 + \mu_t \tag{2.6}$$

While the model in Equations 2.5 and 2.6 has the shortcoming of not testing individual break dates in all parameters, if we are searching for a plausible reason for the obvious volatility break in the mid-1980s it is reasonable to expect that the coefficients might experience a concurrent break. It is possible to construct and test the model for independent breaks of the constant and lag coefficients along with the variance break, but those tests were not performed here. For the U.S. data alone, simultaneous estimation of non-concurrent variance and lag parameter breaks would involve trying to select one model from nearly 30,000 regressions. The MS model is an obvious alternative to searching for the best structural-change model.

2.4.2 Markov-Switching Estimation Methods

Seeing that the purpose of this study is to determine the importance of non-linearities, the MS models have been simply specified. The MS models estimated here have three basic structures. Within each structure for each country we compare models with and without a linear structure. Those models without a linear structure follow an AR(0) specification taking after the work of M&P and Albert and Chib (1993). Work by M&P and Albert and Chib (1993) find that the AR(4) model of Hamilton (1989) is over specified once an expanded U.S. dataset is considered. More importantly, all motion within a model specified as an AR(0) is due to either a current shock or the state of the system. Models are also estimated including a linear function in the form of a simple AR(1) term. Generally, each structural system is specified where $\phi = 0$ or $\phi \neq 0$ and the variable y_t is output. These basic structures will be denoted "NoAR" for Equation 2.7 and "AR(1)" or "linear structure" for Equation 2.8 throughout this paper.

$$\Delta y_t = \mu_{S_t} + \epsilon_t \tag{2.7}$$

Equations 2.7 and 2.8 are assumed to have a normally distributed error term, $\epsilon \sim N(0, \sigma^2)$. Adding an autoregressive term (ϕ) to Equation 2.7 yields:

$$\Delta y_t = \mu_{S_t} + \phi(\Delta y_{t-1} - \mu_{S_{t-1}}) + \epsilon_t \tag{2.8}$$

In both Equation 2.7 and 2.8, S_t is an unobserved random variable that is the realization of a two-state Markov chain. Transitions between the two states (S_i for t and

S_j for $t-1$) in all specifications here are governed by fixed probabilities

$$Pr(S_t = i|S_{t-1} = j, S_{t-2} = k, \ldots, \Delta y_{t-1}, \Delta y_{t-2}, \ldots) =$$
$$Pr(S_t = i|S_{t-1} = j) = p_{ij}$$

The parameters of these models are estimated using maximum likelihood based on a version of the EM algorithm as described by Hamilton (1989). The variance σ^2, and state specific means μ_1 and μ_2 are estimated along with p_{11}, and p_{22} and ϕ. The value p_{11} represents the probability of remaining in state 1 (typically the high-mean growth state) in time t given that you are in state 1 at time $t-1$. Likewise p_{22} is the probability of remaining in state 2 (typically the low-mean growth state), given that you are in state 2 to begin with. The transitional probabilities that are most interesting can be calculated as $p_{21} = 1 - p_{11}$ and $p_{12} = 1 - p_{22}$. Each transitional probability represents moving from the high-mean state to the low-mean state, and vice versa. In this simple structure it is assumed that there is constant volatility.

A second structure is estimated, where the mean and variance are "tied" to one another. The TS structure assumes that each state experiences its own measure of volatility. Thus, recessions have one level of volatility, while expansions have their own volatility. These models are popular methods of accounting for heteroskedasticity across states. Krolzig and Toro (2005) is one of many examples using the tied volatility MS model in practice. The TS variance models are governed by only a single transitional probability like in Equations 2.7 and 2.8. The TS models specified in Equations 2.9 and 2.10 are estimated both with and without a single AR(1) term.

$$\Delta y_t = \mu_{S_t} + \epsilon_{S_t} \tag{2.9}$$

$$\Delta y_t = \mu_{S_t} + \phi(\Delta y_{t-1} - \mu_{S_{t-1}}) + \epsilon_{S_t} \tag{2.10}$$

Equations 2.9 and 2.10 are specified such that the volatility is state specific, and $\epsilon_{S_1} \sim N(0, \sigma^2_{S_1})$ and $\epsilon_{S_2} \sim N(0, \sigma^2_{S_2})$. It should be noted that the volatility shifts modeled in Equations 2.9 and 2.10 are distinctly different than those shifts identified by the Great Moderation. The Great Moderation implied a wholesale break in volatility and not necessarily a change in state-specific volatility.

The third structure estimated allows for the volatility to shift "independent" of the mean. The independent switching or IS variance model was first used to analyze GDP for the U.S. by McConnell and Perez-Quiros (2000), but has only recently been used in fitting business cycles. Hamilton (2005) provides a more detailed description of the inference and estimation methods used here. In this structure there are two independently determined rates of growth for each of two separate levels of volatility. Thus there are four mean growth rates in all. Switching between the high- and low-mean and high- and low-volatility states are each governed by a separate two-state Markov chain.

$$\Delta y_t = \mu_{S_t,V_t} + \phi(\Delta y_{t-1} - \mu_{S_{t-1},V_{t-1}}) + \epsilon_{S_t,V_t} \tag{2.11}$$

$$\Delta y_t = \mu_{S_t,V_t} + \epsilon_{S_t,V_t} \tag{2.12}$$

Volatility states are governed by a similar method as growth states

$$\begin{aligned} Pr(V_t = i | V_{t-1} = j, V_{t-2} = k, \ldots, \Delta y_{t-1}, \Delta y_{t-2}, \ldots) = \\ Pr(V_t = i | V_{t-1} = j) = q_{ij} \end{aligned}$$

The probability of switching from the high-volatility state V_1 to the low-volatility state V_2 is q_{21} and represents a shift much like the Great Moderation. Equations 2.11 and

2.12 represent the most complex models examined here. Means under a high-volatility regime (V_1) will independently switch between a high-growth (μ_{11}) and low-growth (μ_{21}) state. If the model transitions into a low-volatility regime (V_2), the mean will independently switch between μ_{12} and μ_{22} according to transitional probabilities p_{12} and p_{21} respectively. Variances are estimated for the two separate states as a high-volatility state (σ_1^2) and low-volatility state (σ_2^2). Unlike work by Smith and Summers (2005) there is no attempt to constrain the variables to strictly represent recessionary or expansionary states as negative and positive growth respectively. Thus it will be shown that several models predict positive growth during their low-mean growth states. However, these models often coincide with business cycle turning points. Models represented by Equation 2.12 are the best fitting models for business cycle characteristics for each country.[8]

Statistical testing of MS models is complicated in a manner similar to testing for an endogenous break in the volatility of a linear model. The probability of shifting states represents an additional nuisance parameter when testing between structural-change models and basic MS models. Therefore the tests to perform the statistical test between structural-change models and MS models are not standard. Furthermore, there are additional nuisance parameters between CV and TS models, and again between TS and IS models. Therefore, in order to statistically test each of these increasingly complex models, new test statistics must be created through Monte Carlo simulation with the appropriate degrees of freedom, similar to what was carried out in Carrasco (2002) and Hansen (1997). As the purpose of this research is not to select the best fitting model in parametric terms, we continue by examining the commonly

[8]Estimation was performed using programs made publicly available by Hamilton (1989) from his website listed in Appendix A. Independent switching models were estimated using programs made publicly available by CPQ also listed in Appendix A.

used models in the literature and comparing their ability to replicate non-parametric features.

2.5 Estimation Results

The first linear estimations are estimated with a focus on attempting to determine the presence and location of a break in the variance. The results here confirm the findings of the 1984Q1 date for a U.S. volatility break which closely matches estimates by McConnell and Perez-Quiros (2000) and Kim and Nelson (1999). For the U.K. a volatility break was detected in 1992Q2 which roughly agrees with estimates by van Dijk, Osborn, and Sensier (2002). A volatility break was estimated to have occurred in Australia during 1985Q2. Linear models of various ARIMA(p,d,q) specifications could be tested and compared. However, it should be noted that the break tests are selecting a model based upon the most likely period for a parameter shift. Additionally, once a structural change has been detected in the linear model, specification tests like AIC/BIC tests do not confirm that the same linear structure is necessarily optimal in both periods. Thus, it seemed most logical to focus on comparing the same basic models across countries in lieu of estimating a multitude of ARIMA models.[9] Subsequently, the MS estimations are performed simplifying the need to simultaneously model and test non-linear and linear features. Statistical tests of the various MS estimations versus linear models with and without volatility breaks are not performed here. A more complete explanation on the computational difficulties and current testing theory is given in Appendix A.

[9]Specification tests for the ARIMA models are available from the author upon request.

2.5.1 Linear Regression and Endogenous Break Test Results

The estimation of the system of Equations 2.3 and 2.4 for the U.S., the U.K., and Australia all indicate a break in their volatility as can be seen from the test statistics in column 3 of Tables 2.1 - 2.3. Parameter estimates are shown for a random walk with drift model (column 1), a GMM model estimated with a constant volatility (column 2), and a GMM model allowing for a structural break in volatility (column 3). Wald test statistics shown at the bottom of column 3 for the U.S. in Table 2.1 are all highly significant and estimate the most likely break date for the variance at 1984Q1. A similar Lagrange Multiplier test for parameter stability of the U.S. model using p-values corrected using methods suggested by Hansen (1997) also estimates the break date at a nearby 1983Q1, while a Likelihood Ratio test estimates the break date at 1983Q2. Any of these dates could plausibly be used in the forthcoming Monte Carlo estimation, but the Wald test dates were chosen due to the fact that they agree with previous estimates and their ease of calculation. Each of the models in columns 1 - 3 in Tables 2.1 - 2.3 are tested in the simulation section. Carrasco (2002) notes that structural-change tests of a linear model may not be correct if the true model is a threshold autoregression (TAR) or MS model. However, Carrasco (2002) only confirms this result for a model with a shift in mean and AR parameters and does not conduct similar tests with shifting innovation variance. Carrasco's finding that structural-change tests may not be correct relies on recreating test statistics using known models and Monte Carlo simulation. While these findings are important to understanding the model selection process, the problems Carrasco raised were not pursued further here. Therefore, the Likelihood Ratio tests performed here assume standard distributions, since the non-standard distributions are not well understood. It is worth exploring these data for additional breaks to understand where the shortcomings of the simulations might be even though the point of this research is not to find an optimal linear model.

It can be seen in column 3 of Table 2.1 that U.S. growth volatility moderated from approximately 1.0996 to 0.4781 between the pre- and post-break periods. The coefficients reported here are based off of quarterly changes in output measured as 100*Δln(output). Thus, a constant value of 0.5118 for the U.S. in column 3, translates into approximately 2% annual growth ignoring the persistence of the AR(1) term. The U.K. experienced an even greater decline in volatility falling from 1.0565 to 0.3779 at approximately the second quarter of 1992. For the U.K. it should be noted that the break date of 1992Q2 was used for simulations, but a Lagrange Multiplier test estimated the U.K. break date at 1985Q2. Australia also experienced a significant decline in volatility around 1985Q2, as the level of volatility fell from 1.3994 to 0.6111.

Tables 2.1 - 2.3 also display tests for subsequent volatility breaks in the remaining shortened periods, conditional upon the first break. Each subsequent break test was performed on only the central 70% of the remaining time series to avoid problems with endpoint break detection. The coefficients reported in columns 4 and 5 in Tables 2.1 - 2.3 are from the period in which the Wald statistic predicts the greatest probability of a break in the standard deviation. For the U.S., there is some evidence of an additional break around 1987Q3 in the second half of the data with *sup* tests exhibiting low p-values. However, when the *ave* and *exp* tests are examined, there is insufficient evidence to declare an additional break for U.S. data. The U.K. estimates in columns 4 and 5 of Table 2.2 show little evidence of any additional breaks. However, tests of additional breaks in Australian GDP volatility show substantial evidence of an additional downward volatility break in 2004Q1. This possibly important finding in the Australian data merits much further study and will be explored in future work.

Given that this volatility shift in Australia has only been present for about four years we believe that furter reductions in volatility would only result in a better fit for the structural-change model. It is possible that output volatility in Australia has simply been trending downward, or there may be more volatility periods than first believed. It should also be noted that the U.K. and Australian models had insignificant autoregressive coefficients in the primary estimation (column 3), particularly in the period prior to the second break for Australia as seen by comparing columns 3, 4 & 5 of Table 2.3. While these anomalies will have some impact on the simulated data and the ability to replicate the features of each country's output, one of the major goals of this research was to see if commonly used simple models of output could successfully mimic business cycle features for any country. These results serve as mounting evidence that there is not one simple model of output that can be applied broadly.

Plots of the Wald test values for each of the possible break dates (excluding the beginning 15% and ending 15%) suggest that U.S. growth had not finished "breaking" by the late 1980's as seen in Figures 2.1 - 2.3. The vertical axes in Figures 2.1 - 2.9 are the *sup* Wald test values, and need to be carefully interpreted. In order to declare a break, all test values (*sup*, *ave*, *exp*) should be significant at some standard level. The period where maximum *sup* Wald test values are obtained are reported in Tables 2.1 - 2.3. These non-standard Wald test values, have their p-values calculated according to the methods laid out in Hansen (1997), and depend on the sample size which if varying in each period.

The *sup* test statistics for the detection of additional breaks in Australian volatility after the initial moderation show some evidence of a second downward shift. Examination of Figure 2.9 shows that there is mounting evidence of a second volatility break where the *sup* test statistic values exceeding 50, yielding a p-value of 0.000. Using all test statsitics (*sup*, *ave*, *exp*), one might conclude that there is an additional break in Australia. Several researchers, including CPQ test a permanent one-time switch in volatility at the pre-estimated break date using dummy variables. While a three-state MS model is not used here for Australia, our estimates and simulations when using MS modeling allow for volatility to switch into and out of high- and low-volatility states.

Estimates for the models estimating Equations 2.5 and 2.6 with simultaneous breaks in all coefficients are shown in Appendix Tables A-1 to A-3. The tests in Appendix Tables A-1 to A-3 examine the possibility that the behavior of the model might be due to shifts in either the constant, standard deviation, or AR parameters. For the U.S. there is virtually no evidence suggesting that the behavior of an ARIMA model might be improved if simultaneous breaks were to occur in either the constant or AR coefficient. Out of the few models estimated here, the ARIMA(1,1,0) model with a variance break was determined to be the best fit.

For the U.K., there is evidence of an AR break if the standard deviation and constant are concurrently allowed to shift as can be seen in Appendix Table A-2 column 2. The maximal point of the Wald statistic for a break in the AR coefficient occurs during 1999Q3 however, and not closer to the time of the Great Moderation. There is also some evidence that the constant is simultaneously breaking. A final test is conducted on U.K. data where the AR coefficient is tested to break, at the same time as the standard deviation is shifting as seen in column 4. However, the *exp* tests do not agree with the *sup* and *ave* tests for the simultaneous break of these models.[10]

Australian regression results in Appendix Table A-3 provide evidence in favor of

[10]See Appendix Table A-2, columns 1 and 7. Given that most of the models estimated here for the U.K. have difficulty simulating stylized business cycle facts, it might be worth exploring breaks in the AR coefficient for the U.K. in more detail.

an AR(1) break. The evidence for the AR(1) break does not hold however if the constant is not also allowed to break as seen in column 7. Column 8 examines the AR(1) break if the SD and constant are not allowed to simultaneously shift. Evidence shows that this estimation might be valid, but given the evidence on the occurrence of Great Moderation this specification is not pursued further here. In the U.K. and Australia it appears shocks have become more persistent after the Great Moderation in the structural-change models. In future work it would be interesting to examine the validity of the AR switching models for the U.K. and Australia.

2.5.2 Markov-Switching Results

Markov-switching models were estimated for each country using models outlined by Equations 2.8 - 2.12. The MS models for all countries with each of the six combined structures converge.[11] Given that numerous researchers have noted that MS models often find local maxima, a variety of starting points were tried for each model estimated here. Parameter estimates for the models tested here predominantly hold given different starting points, however a grid search for optimal starting points was not performed. Those starting values which yielded strong convergence and the highest likelihood were used. However, there is no guarantee that the MS models estimated here are not a local maximum.

The results for the MS models statistically select those including volatility switching like the MS tied switching (MS-TS) and lean towards independent variance switching (MS-IS) models for each country via their log likelihood values. As is previously noted, formal Likelihood Ratio tests across MS models were not performed due to the

[11]Recall that there are three different volatility structures to be compared, each with and without a linear AR(1) structure.

fact that the test statistics have nonstandard distributions and are not well understood. The conclusion of an improved fit is derived from a standard LR test statistic. Given that there is only one additional parameter between a CV-NoAR model and a TS-NoAR model the χ^2_n test has one degree of freedom so the critical value would be approximately 4.0 under normal conditions at the 5% significance level. MS model selection against linear models can be performed using methods proposed by Garcia (1998). Garcia (1998) does not provide any test statistics for comparing a MS model to a structural change in volatility model.[12] In many cases very large LR test statistics (where the Log L increases by more than five units) between models is taken to be highly suggestive of an improvement in fit. Standard LR test statistics can be used when the models being compared have the same basic form and do not have an additional nuisance parameter such as adding a linear structure to a CV model. When comparing the CV versus TS or TS versus IS models the tests are not standard. It can be seen in Table 2.4 that TS and IS models show dramatic increases in likelihood values over CV models for each country.

The U.S. TS variance models that are popular in other research show findings that square with much of the previous literature. The high-mean growth state is associated with lower volatility in the TS-NoAR model, and low-mean growth states have higher volatility. However, the low-mean growth rate for TS-NoAR is insignificant. When including an AR coefficient, as seen in the TS-AR(1), the higher volatility state is actually associated with higher growth, and the AR term is highly significant. It can be seen that there is a significant increase in fit between the TS-NoAR and TS-AR(1) models, with a standard LR test selecting the latter model as a better fit. It is unlikely that a statistical test would select the IS model over the TS model for the U.S. as

[12] See Appendix A for more information on testing difficulties.

there is only a marginal improvement in the log likelihood value when allowing for the volatility to switch states independently of the mean. The MS-IS model for the U.S. shows that the high-volatility high-growth mean state (μ_{11}) is 1.374, while the low-growth state under a high-volatility regime (μ_{21}) is -0.148. Once the U.S. switches into a low-volatility state (which occurs with less than 1% probability ($1-q_{11} = .006$)) the average high-growth state is a slower 0.903 (μ_{12}) and the low-growth state (μ_{22}) is a positive 0.192. These estimates agree with the general findings of other researchers like CPQ and Kim and Nelson (1999) who find that mean high- and low-growth rates have partially converged. However, these findings show that the MS model does not necessarily switch into a "recession" in a low-mean growth state. Examination of the predictions of these MS models yields yet more insight on which model might be preferred.

The MS models produce both filtered and smoothed probabilities showing the likelihood of being in a given state during a given period. All plots produced here showing the filtered probability of being in a certain state use all information up to time t. Smoothed probabilities are produced using all available information and can be used to rule out "false positives" which frequently occur in filtered probabilities.[13] Using the assumptions of Hamilton (1989) economies are predicted to switch states when a probability of being in a certain state exceeds 0.5.

Examining Figure 2.10 shows that the simplest model (CV-NoAR) nearly predicts each recession for the U.S. selected by the BBQ algorithm as shown by the shaded regions in each figure. The MS-CV-NoAR model is even able to predict the 2001 recession that the BBQ algorithm misses. This simple model is also able to predict

[13]Plots of both smoothed and filtered probabilities are available for each country and model from the author upon request.

that the economy entered a recession in 2008. Given that the MS-CV-NoAR model is almost certainly not statistically optimal and researchers tend to account for switching variance, other options are analyzed. Figure 2.13 shows that the tied-switching model (MS-TS-AR(1)) is unable to predict recessions despite its statistical superiority to the MS-CV model. Instead, the tied switching model including an AR parameter is dominated by the change in volatility. As soon as the Great Moderation arrives the probability of the low volatility, high mean growth state is nearly 1.00.

By untying the mean and variance, the MS-IS-NoAR model can improve predictions even further. Around the first period of 1984, the probability of being in the low-volatility state (left-hand vertical axis) moves to near 1.00. The prediction of being in the low-volatility state after 1984Q1 agrees almost exactly with the existing evidence on the Great Moderation. The MS-IS-NoAR model shown in Figure 2.14 displays that the volatility switch is detected by a dramatic shift in the probability of a low-volatility state. This MS-IS-NoAR model also selects every slow-growth or recessionary period for the United States. However, the filtered probabilities would lead one to make a false positive prediction in some cases if the $p_{21} > 0.5$ rule is used. Smoothed probabilities in Figure 2.16 for the MS-IS-NoAR model show that there was a high likelihood of entering a recession in 2007. It is worth noting that the probability plots for the MS-IS-AR(1) model with linear features are essentially indistinguishable from the model without an AR term. The AR term in the MS-IS-AR(1) model for the U.S. is insignificant at the 5% level, and appears to add little information. A standard Likelihood Ratio test between MS-IS-NoAR and MS-IS-AR(1) models conclude that the linear structure does not significantly improve the fit of the model at the 5% level.

The most innovative finding from examining the probability plots of the MS-IS models, is the recent decline in the probability of being in the low-volatility state. Beginning in 2008, the probability of being in the low volatility state was no longer very close to one. There is not enough evidence at this time to show that the Great Moderation is over, but there is enough evidence to consider this a real possibility.

Estimates for the U.K. lend even more evidence that the Great Moderation might be ending on a global scale. Parameter estimates for the same six non-linear models examined for the U.K. are displayed in Table 2.4. Examination of the probability plots of the MS-CV-AR(1) model for the U.K. in Figure 2.17 shows that a simple MS model can predict many but not all of the U.K. business cycle turning points. Removal of the AR term from the MS-CV model leads to similar state probabilities. The MS-CV-AR(1) model is shown here due to the significance of the AR term in the regression. Tied-volatility models for the U.K., lead to similar state probabilities as for the U.S., where volatility switching dominates the states as seen in Figure 2.13. Log likelihood values would almost certainly select a switching variance model over a homoskedastic model. Likelihood Ratio tests between TS and IS models lean towards selecting the IS models over others. At standard test levels the IS-NoAR would be preferred over the TS-NoAR, but due to the non-standard test statistics formal tests were not performed.

In the U.K. MS-IS-NoAR model, the high-volatility, high-mean state has a quarterly growth rate of approximately 0.762%. The probability of remaining in this state is $p_{11} = 0.955$, implying that the high growth state is relatively persistent. Also, the low-volatility state, σ_2^2, has an exceptionally low measure of 0.065. Also notable about the low-volatility state, is that the slow-growth state, μ_{22}, is predicted to be an average of 0.615, not exactly what most economists would refer to as a recession. The plot of filtered probabilities in Figure 2.18 can display this result more explicitly.

According to these predictions, the U.K. has very likely been in the low-mean growth state on and off since about 2001. This shift into the low-mean growth state includes the recent shift out of the low-volatility state. Note that in the high volatility state it is still likely the U.K. economy remains in the low-growth state where the average quarterly growth rate would be -0.057. Examining Figure 2.19 shows that the probability the U.K. is in the low-volatility state is now close to zero. During the period of the Great Moderation the U.K. did not really experience any true recessions, but rather a long period of high growth and an ensuing period of slower growth. The MS-IS-AR(1) model is not statistically preferred to the NoAR model according to standard Likelihood Ratio tests, and the AR coefficient (ϕ_1) is insignificant.

Australia's MS-IS model estimates are quite different than those predicted for the U.S. and U.K. with regard to volatility. As was noted when measuring the linear regressions, it is possible that there were breaks in the AR or constant parameters. In the TS and IS specification estimates for Australia, the AR coefficients are insignificant. Despite this fact, these models were estimated here in order to compare the results across typically used methodologies. Others such as Kim, Morley, and Piger (2005) have used the models tested here to estimate business cycle turning points for Australia. Likelihood Ratio tests would lean towards selecting the MS-IS model over other alternatives. Figure 2.21 shows that the CV model with an AR coefficient is poor at predicting recessionary states. The model should predict a total of 7 peaks to square with the BBQ algorithm, but actually only estimates four low-mean states. An examination of Figure 2.22 shows the MS-TS-AR(1) model is dominated by the volatility shift in 1985, as was seen in the U.S. and U.K. probability plots. Finally, Figures 2.23 and 2.24 show probabilities for the independent variance switching models.

The IS model without a linear structure is superior statistically according to standard Likelihood Ratio tests, however neither of the IS models happens to do a good job predicting recessionary or low-growth states. Both Figures 2.23 and 2.24 predict a low-growth state in 2001, that is not captured by the BBQ algorithm. Also of note, the period prior to the Great Moderation appears to have little ability to select turning points in the data. It would appear that none of these parametric models are really able to fit the turning points that are selected using the BBQ algorithm. The problems pointed out during the linear estimation phase may play a much more important role in the Australian model, including the possibility that there was an AR switch at some point. The low-volatility state in Australia has a higher separation in the average high- and low-growth states. As seen in Figures 2.23 and 2.24 it is possible the MS model is only appropriate following the Great Moderation. Taking a closer look at the simulations that these models produce will yield yet more insight as to how these economies should be modeled.

2.6 Simulation Procedures

The parameters that are estimated during the previous sections, are used to simulate data to compare to the original GDP data for each country. Each simulation is estimated assuming there is no parameter uncertainty for any model. Errors are assumed to be distributed normally using the estimated standard deviations from the estimation phase. Alternative estimates using bootstrapped residuals yeilded similar results are not reported here. All simulations are produced to have the same number of observations that exist in the underlying data. Every model is simulated 5,000 times, and then compared to the original data as well as simulations produced by other models. Table 2.5 shows the various simulation experiments conducted. For the linear models with a structural break in volatility, additional simulations are conducted with modifications to volatility. Models 3 - 6 test variations of the linear model, where the volatility is held constant at the first (and second) period level for the entire time period. For the U.S. this amounts to running a 61-year simulation with high (and low) volatility, when in fact only 37 years were high volatility. Additional simulations were conducted where volatility switches at the date predicted by the Wald tests performed in Tables 2.1 - 2.3. A final counter-factual linear simulation is conducted where the high volatility is in place for the shorter timespan ("SD Reversing"). The reversing volatility estimates, would help to show if the ARIMA(1,1,0) models would produce appreciably different simulations if low-volatility were to persist for several more years. Model 5, with switching volatility at the date of the Great Moderation, is the model that should be used as the primary comparison to the MS simulations. Model 2, a simple linear model, is typically compared to the MS approach in other studies. The six MS models estimated for each country are simulated here. Models 11 and 12 are the IS models that tend to fit the data the best for each country. Unlike studies by CPQ, KMP, and M&P, the MS volatility switching is allowed to occur according to a random draw rather than by using a dummy variable fixing the volatility switch at a preassigned date.

2.6.1 Q-Test and Kolmogorov-Smirnov Test Statistics

After *each* of the Monte Carlo paths are simulated, summary statistics are calculated using the dating algorithms and stored for each run. The simulation statistics are then collected into empirical distributions to compare the values of the stylized facts to the distribution of simulation statistics. Reference characteristics are calculated using the same algorithm used to measure the simulations. This method is a departure from previous papers such as CPQ who test algorithm measured features against the NBER averages, or M&P who use an optimized algorithm. In addition to examining the empirical distribution functions from the simulated data, statistical tests are performed to see if the simulated data is a reasonable approximation to the actual underlying series. In many of the simulated series, no turning points are detected over the entire length of the data. These simulations are simply dropped from the analysis. For the most part, the analysis does not change if these zeros are included. Unreported empirical distributions show similar conclusions to the ones reported here.

As noted by Hess and Iwata (1997), the joint distribution of the business cycle features is unknown. Using a generalized Q-statistic allows us to test the features of the simulated distributions with the underlying distributions. The generalized Q-statistic was used by Hess and Iwata (1997) and Simkins (1994) when studying business cycle characteristics like those studied here. Cogley and Nason (1995) also used Q-statistics to test the lags of impulse response functions and autocorrelation functions of simulated data. The Q-statistic used here takes the form:

$$Q = (\hat{B} - B)'\hat{V_B}(\hat{B} - B) \tag{2.13}$$

$$\hat{V_B} = N^{-1}\sum_{k=1}^{n}(B_k - B)(B_k - B)' \tag{2.14}$$

$$B = N^{-1}\sum_{k=1}^{n} B_k \tag{2.15}$$

where $\hat{B}$ represents a vector of business cycle characteristics estimated from the simulated data, while B represents the same average features listed at the bottom of Table 1.2 for each country. The value of $\hat{V_B}$ is the covariance matrix of the characteristics being tested, and is estimated using Equation 2.14. Each value of B_k represents the average for the feature of the business cycle from the k^{th} simulation. Thus the Q-statistic is essentially examining the aggregate distance of simulations features away from the features underlying the estimation. The test statistic is distributed approximately $\chi^2(n)$ with degrees of freedom n equal to the number of business cycle features being tested. Hess and Iwata (1997) note that the Q-statistic might be unreliable as it assumes the distribution of $\hat{B}$ is asymptotically normal.

In addition to the Q-statistics, empirical distributions are examined relative to the values of the underlying data series features. For each of the features measured here a 90% confidence interval is examined. If the actual value is in either the upper or lower 5% tail of the empirical distribution, the simulated feature is deemed to not fit. Clements and Krolzig (2004) note that empirical distributions with high variability will have a high probability of capturing the true value of the distribution. Thus, flat empricial probability distributions may lead to decieving Q-statistics. The empirical distribution statistics can also be useful if one were inclined to use the NBER averages rather than the BBQ averages. If these NBER figures fall into the 90% distribution as well, one could then conclude that the empirical distributions were not appreciably different.

In order to compare the models against one another, Kolmogorov-Smirnov ("KS") test statistics are conducted for each feature. Hess and Iwata (1997) only test for contraction and expansion features, and note that KS-tests are typically too conservative in nature. The major drawback of the KS-test is that it can only test one type of feature at a given time, rather than a joint test. However, the KS-test is particularly useful when comparing one simulated model against another. The KS-tests are setup such that $F_n(x)$ is the empirical distribution and $H_n(x)$ is the cumulative density function for any particular feature from another model. Each feature can be tested to see if they are likely from the same distribution. The statistic $D^+ = sup_x(F_n(x) - H_n(x))$ is used to determine if $F_n(x)$ is significantly different from $H_n(x)$. In almost all cases, the simulations report statistically different distributions for features. However, these tests and figures can easily display the differences between two simulations.

2.6.2 Simulation Results

Tables 2.6 - 2.8 show for each country, the ability for a given model to simulate the specified features of actual GDP. The columns show separate tests for the ability to jointly estimate certain features of the model. The first column shows a test of all eight features of the model simultaneously. The eight features are duration, amplitude, contraction, and excess, testing for both expansions and contractions. Columns 2 and 3 test recession and expansion features. Pairs of features are also tested jointly in columns 4 through 7. For each feature, the expansion and contraction simulation characteristics are tested. Thus, these tests can see if a model can successfully simulate both expansion and contraction features of the data such as both duration measures as measured using the BBQ algorithm. The last two columns of Tables 2.6 - 2.8 show a pair of tests on the business cycle features of "duration and amplitudes" and "cumulative and excess" together. The final two columns test for four characteristics jointly. For example, "durations and amplitudes" tests both the duration and amplitude measures jointly for recessions and expansions. The "cumulative and excess" category is a basic test for the ability to simulate the asymmetric features of the underlying data. In Tables 2.6 - 2.8 a value of zero or close to zero indicates that the characteristics from the simulation are statistically different than the underlying feature. Values greater than 0.10 fail to reject a difference in the the simulations and underlying features from Table 1.2.

Table 2.6 shows that the only models that successfully mimic all characteristics of the underlying U.S. GDP data are produced using simple structural-change in volatility models and the independent switching version of the MS model. The MS-IS-NoAR and AR models for the U.S. successfully fit all characteristics of the underlying data. However, recalling the fact that the AR coefficient is insignificant in the estimation phase helps to justify the use of the MS-IS-NoAR model. Both the TS and CV versions of the MS model perform very poorly, but notably do a decent job replicating recessionary characteristics. Only the simplest MS-IS models are able to simulate recessionary features with any accuracy. It can be seen clearly here that linear features do not help improve the fit of any of the MS models for the United States. The structural-change models for the U.S. create reasonable fitting simulations, but are notably close to creating significantly different measures of excess relative to the underlying data. The linear Great Moderation model ("SD Switching"), performs well enough that it is difficult to rule this model out. These tests of the U.S. simulations are evidence that the Great Moderation must be properly accounted for in estimation, either using an MS model or by using a structural-change in volatility model. The simple homoskedastic linear specification is not able to match many of the stylized facts, and performs significantly worse than the structural-change model.

Tables 2.9 and 2.10 present statistics from the empirical distributions for the the linear and MS models respectively. For each country, the actual values for characteristics are reported in the first column, and the average of the simulations is presented for each model. The average value for each stylized fact for both contractionary and expansionary features are then compared to the empirical distributions by calculating the percentile of observed simulation statistics that are less than the value for the stylized fact. Bold percentages are those where the actual stylized fact falls in the center 90% of the empirical distribution. For example, in Table 2.10, under the CV-NoAR header, for peak-to-trough (P→T) durations, the average length of a recession in simulation was 3.92 quarters, and the stylized fact value of 2.889 falls into the 3rd percentile of the empirical distribution. This implies that the CV-NoAR model for the U.S. is unable to create recessions that are short enough to match the stylized facts. Ideally we would like to see values of 50% for each of the empirical distributions, meaning that the simulations produce a given stylized feature in one-half of all simulations. However, in trying to narrow down the best model for estimation it is worth knowing which models are able to replicate business cycle features with some frequency.

In support of the Q-tests shown in Table 2.6, empirical distributions in Table 2.9 display that the linear and switching models have modest success replicating the underlying cyclical features in U.S. GDP. However, as many others have found, linear models have tremendous difficulty estimating the asymmetry present in U.S. business cycle data, as seen in the CV-NoAR model where the T→P empirical distribution fails to emulate anything as large as the stylized excess measure in 100% of simulations. The IS-NoAR model in Table 2.10 for the U.S. is able to match asymmetric features providing some evidence that the "bounceback" model promoted by KMP has merit as it is specifically designed to capture this feature. Also, Table 2.10 shows that the independent switching variance models are excellent at simulating the actual business cycle features. Each of the MS models that are not IS volatility have trouble replicating the expansionary features of data. This reveals the importance of testing these models for both expansionary and contractionary features.

For the U.K., linear models have mixed success. Shown in Table 2.7, both the switching and reversing volatility models are able to replicate most expansionary features of the business cycle, however no linear model is able to capture all the recessionary characteristics. Tables 2.9 and 2.10 show that no linear or non-linear model is able to match all business cycle features individually in the U.K. simulations. The MS-IS-NoAR model previously shown to be a good simulator in Q-tests, fails to replicate the amplitude and cumulation features of cyclical contractions in at least 5% of simulations. The tendency for the MS models for the U.K. is to underestimate the depth or magnitude of recessions. Similar to the findings for the U.S., the TS and IS models are able to replicate most features. However, none of the MS models are able to capture all the features that are studied here. For the U.K. the TS model is able to replicate the contractionary features, but fails to replicate expansion features. The IS model performs better in the expansionary periods, but like the linear models predicts recessions that are too short and shallow. In the U.K. the tied-switching model appears to be the best representation of the data. As was previously mentioned, further study with shifts in the AR parameter might be in order. These results provide further evidence that one cannot only examine expansionary characteristics, and must consider the expansionary characteristics as well.

Tests of Australian simulations reveals results very different than those seen in the U.S. and U.K. data. Table 2.8 shows random walk and structural-change models are able to simulate all of the features of the data. Australia is the only one of the three countries studied where a pure linear model is able to replicate all business cycle features jointly. Linear models with shifting volatility are also able to replicate all business cycle features. Additionally, the independent volatility switching model (MS-IS-NoAR) is also able to replicate almost all business cycle features for Australia. This is substantial evidence that an MS model might be able to replicate underlying business cycle features for Australia. For Australia, the models previously shown to simulate the data well jointly, are also very good at replicating individual features. In fact, four of the linear models, including a random walk, are able to outperform all but one of the MS Models that were estimated. As seen in Table 2.10, the IS-NoAR is the only MS model capable of fitting all business cycle features. The inclusion of an AR term appears to hurt the ability of the IS model when simulating data.

In summary, the estimation and simulation results for the U.S. lead us to select the MS-IS-NoAR model as the preferred model. Parametric measures of fit, while not definitive due to non-standard test statistics, lead us to believe that the MS-IS-NoAR model is a very parsimonious model that fits the data very well. Also, the state probability plots provide evidence that the MS-IS-NoAR model is able to select the correct recessionary periods in both high- and low-volatility periods given only GDP data. Finally, the MS-IS-NoAR simulated data are able to mimic nearly all of the business cycle features seen in actual U.S. GDP data. Overwhelmingly, the results here for the U.S. lead us to believe that a homoskedastic or tied-switching model is an inadequate representation of the economy. In order to reach this conclusion, it should be noted that the structural-change in volatility is the proper model for comparison and not a plain linear model.

For the U.K., parametric models lean towards selecting the MS-IS-NoAR model, as do probability plots and Q-test statistics. However, the empirical distributions for the U.K. show evidence that the MS-IS-NoAR model does not reproduce recessions that capture the depth of actual recessions. These findings lead us to believe that the MS-IS-NoAR specification might be a good starting point for augmenting a model designed to capture the larger downturns seen in U.K. data.

Australian parametric tests support the MS-IS-NoAR model. However, Q-tests of simulated data for Australian models fails to rule out linear or structural-change in volatility specifications. Probability plots show that the IS-NoAR models are unable to pinpoint recessionary periods until the period of the Great Moderation. If we were relying on these probability plots to select a model, the CV-AR(1) specification would appear to be a better match to the recessionary periods signaled by the standard BBQ algorithm. It is not clear after these tests that there is a model for Australia that would be unanimously preferred.

2.6.3 Comparing Different Models

Within each country, Kolmogorov-Smirnov tests compare individual model features to determine if they statistically predict different distributions. For the U.S. the volatility switching ARIMA model performs as well as both MS models in matching simulation characteristics. Kolmogorov-Smirnov tests show that the ARIMA model with a break in volatility (SD Switching) predicts very similar recessionary characteristics as the MS model with independent switching (IS-NoAR) as shown in Appendix Figure A-1.[14] Comparison of the expansionary features for the U.S. between the SD Switching and IS-NoAR models show that the ARIMA model predicts significantly shorter average expansions and more linear contractions. In general, by examining Tables 2.9 and 2.10 it can be seen that the ARIMA model has much closer predictions to the "true" characteristics determined by the BBQ algorithm. Examining the lower right panel of Appendix Table A-2 shows that the MS-IS-NoAR model has much wider dispersion of the excess feature, implying that it is able to create excess in many simulations. If, as expected, the U.S. recession that began in 2007Q4 lasts

[14]All KS-tests are estimating significant differences at the 10% level in a two-sided test. The KS statistics are not reported here, but are available upon request.

for approximately two years (8 quarters), then the stylized facts for average length and depth of recessions would increase. However, if we offset this increase with the short 2001 recession which was not detected by the algorithm, it is likely that the average recessionary features would not change appreciably. Increased values for the average recession length would serve to increase the fit of the simulations from the SD switching and IS-NoAR models relative to the current simulation. Likewise, expansionary features are slightly overstated in the stylized facts since the most recent expansion as detected by the BBQ algorithm is overstated to a large degree.

Similar distributional KS-tests are performed for the U.K. and Australia. In the U.K. and Australia the empirical distributions for recessionary features of IS models show that linear structures lead to longer and deeper predicted recessions and shorter and less extreme expansionary periods. The KS-tests show that the IS variance switching model does add to the predicted expansion and recession durations relative to the SD switching model.

2.6.4 Autocorrelation & Jump and Rest Tests

A final test was performed for each country, examining the autocorrelations created by the simulations themselves. In support of the jump-and-rest theory of business cycles, the MS models without linear features are able to simulate data for the U.S. that has a positive and significant autocorrelation of 0.2336 on average. A positive autocorrelation for the U.S. is in agreement with CPQ who find that these models which have no persistence built in are able to create data that has a positive autocorrelation. For the U.K., a positive 0.1074 autocorrelation was predicted by the MS-IS-NoAR model. Australian simulations of the same non-linear model were able to produce autocorrelations of 0.510, which were typically highly significant. For the

U.K. and Australia, the presence of a positive autocorrelation is not supportive of the observed data. This mixed evidence provides skepticism on the jump-and-rest nature of business cycles across economies using MS models with independent switching behavior.

Tests for the autocorrelation after accounting for recessionary periods as in Equation (2.2) show that if growth is regressed on a constant, and lag there is a positive 0.3367 lag coefficient. By including the NBER recession dummy variable as in CPQ, the autocorrelation falls to an insignificant 0.101. The same regression run on a BBQ recessionary state dummy variable yields a significant autocorrelation coefficient of 0.1548 with a t-statistic of 2.804. The estimated autocorrelation using the BBQ model recessionary states is much closer to insignificant, but cannot fully control for growth. Camacho and Perez-Quiros (2006) used the BBQ algorithm in a similar regression of U.S. 1980 - 2004 data finding that the autocorrelation was insignificant at the 10% level. Using the MS-IS-NoAR model, a "recession" dummy variable was created where the state predictions from the smoothed probability of a low-mean state exceeded 0.5. The MS-IS-NoAR turning points predicted an insignificant 0.0792 autocorrelation coefficient. This provides some additional evidence that the MS model with independent switching and no linear structure is able to show that U.S. growth behaves according to the jump-and-rest hypothesis of CPQ. For each of the recession dummy variable regressions run here for each country p-values from Ljung-Box and Durbin-Watson tests of the residuals rejected the hypothesis of autocorrelation. These results provide evidence that MS models are able better to predict NBER turning points properly even though the states are not fully coincident.

Autocorrelation tests of regressions of U.K. growth on a constant and lag yield an insignificant AR coefficient of -0.055.[15] Recessionary states for the U.K. were selected using both the BBQ algorithm and the periods where the MS-IS-NoAR low-mean growth probability exceeded 0.5. In each case, the regression of growth on a constant, lag, and recession dummy yielded a significant value of -0.22. A final test was conducted using recessions declared according to the state predictions of the MS-CV-AR(1) model the lag term coefficient is found to be negative and significant.

For Australia autocorrelation tests of growth on a constant and lag estimates an insignificant lag coefficient of -0.0435. When recession dummy variables are created using the BBQ algorithm or MS-IS-NoAR states, a negative and significant coefficient of -0.19 is estimated. When recessionary states are determined using the CV-AR(1) model the autocorrelation coefficient is estimated as a negative and significant -0.1402. These results for the U.K. and Australia provide some evidence that there is not jump-and-rest behavior in those economies. There is no NBER to determine business cycle turning points for the U.K. and Australia, whose methods might be able to more accurately detect the actual turning points in the economy. While it is apparent that each of these countries experienced a Great Moderation, they all appear to need significantly different models to capture their business cycle features and make more accurate predictions.

2.7 Conclusion

If the Great Moderation is ending, then researchers must rethink many of the causes for more moderate volatility in the first place. Theories of the Great Moderation that relied on technology increasing the ability for firms to better predict the future and

[15] Note that these values differ from Table 2.2 as they are estimated using OLS rather than GMM with two equations.

adjust their inventories more accurately might be misleading, considering that this technology has not disappeared. In all likelihood, the results presented here lend more evidence to the good luck hypothesis, as monetary policy has not changed it's focus substantially prior to the new period of increased volatility. Other theories of the Great Moderation like financial innovation and global integration leading to reduced risk appear to be at odds with the evidence shown here.

Substantial new evidence presented here displays that the Markov-switching model for business cycles is a very good tool for predicting the states of the economy, provided that the growth state and volatility state are free to move independently of one another. The tied-switching model that is used by many should be avoided in practice if the goal of the model is to predict business cycles in GDP. Further evidence shows that MS models are not generally improved through the inclusion of a linear structure. This evidence shows that linear structures within any MS model should be carefully considered if the goal is replication of business cycle characteristics. Statistically, models that contain linear structures and/or tied switching features are acceptable, and sometimes preferable. However, these same models fail at replicating business cycles in the underlying data in many regards. These same TS models are unable to make useful recessionary state predictions as their performance is dominated by the presence of the Great Moderation.

This paper also provides evidence that the MS model is comparable to the NBER business cycle turning points for U.S. GDP. By employing the predicted recessionary states as dummy variables in simple growth regressions for the U.S., it is found that the autoregressive coefficient on output growth becomes insignificant. These results support the notion that U.S. growth is best modeled using a Markov-switching approach with independent switching volatility and no linear structure, behaving according to the jump-and-rest pattern described by Camacho and Perez-Quiros (2007). Additionally, it is shown that U.K. and Australian output growth does not appear to exhibit the jump-and-rest feature. These findings are important evidence that shocks and recessions have vastly different effects in different economies. The varying models preferred across countries might explain why the synchronization of business cycles has been difficult to pin down.

The MS model is not without its drawbacks however. For each country, a structural-change in volatility model accounting for the Great Moderation preforms reasonably well when simulating business cycle features. Thus, it can be concluded that the appropriate comparison for any MS model is a linear model that includes heteroskedasticity. To date studies testing the parametric fit of MS models including heteroskedasticity have supported the use of their models by making their comparisons to linear models that we have shown to have very poor predictive power. Additionally, the U.K. and Australian business cycles are not easily simulated using the MS approach. Empirical distributions show that a MS model including a “bounceback” component might be appropriate to model growth patterns, but that volatility must be allowed to switch independent of the mean growth rate. The fact that structural-change models perform as well or better than the MS model in Australia leaves the door open for linear models to have an appropriate place in business cycle research. Regardless, one should be careful to examine both expansionary and contractionary features when trying to decide which model is appropriate for a certain economy.

If more evidence can be shown that the business cycles behave according to a jump-and-rest pattern, it will be necessary to have a better understanding on why MS models explain the autocorrelation in the GDP growth. Additionally, it is necessary to improve upon the method of measuring business patterns with an algorithm.

The methods that are currently used to determine business cycle turning points using an algorithm are sub-optimal, and may cause some models to be falsely rejected. Ultimately, the Markov-switching model, when used in conjunction with current simple business cycle algorithms help select models that are able to account for most business cycle characteristics.

Table 2.1: Estimated Rates of Growth and Volatility for the U.S.

			GMM Models		
				Test for Additional Breaks	
	Random Walk	No Break	Volatility Break	Period 1	Period 2
	(1)	(2)	(3)	(4)	(5)
Constant	0.8095	0.5354	0.5118	0.5701	0.5304
(P-values in Parentheses)	(0.062)	(0.000)	(0.000)	(0.000)	(0.000)
AR(1)		0.3256	0.3226	0.3343	0.2555
		(0.000)	(0.000)	(0.000)	(0.007)
Std. Deviation (Pre-Break)	0.9620	0.8280	1.0996	1.0137	0.2732
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)			0.4781	1.2740	0.4949
			(0.000)	(0.000)	(0.000)
Wald Test Values					
Sup			60.5030	3.7424	18.4382
			(0.000)	(0.738)	(0.002)
Ave			26.5158	0.5626	4.9563
			(0.000)	(0.798)	(0.008)
Exp			18.7328	0.8789	1.1212
			(0.000)	(0.825)	(0.695)
J-Statistic			2.2073	1.6968	1.9789
Break Dates					
Wald Statistic			1984Q1	1978Q1	1987Q3

Values are calculated for these parameters when allowing for breaks in the specific parameter. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the mean, represents the mean before the break. The second reported value for each parameter when allowing for breaks estimates the value after the break.
*Wald tests here signify a test of the restriction rather than a test for the break date. The parameter values are those from the estimated break date supported by the model.
Column (1) represents a random walk model of growth, and estimates the standard deviation of the residual values.
Column (2) estimates a basic ARIMA(1,1,0) model of output using GMM Equations (1) and (2) with no break.
Column (3) estimates an ARIMA(1,1,0) model of output with parameter breaks std. deviation.
Columns (4) and (5) estimate the same model as column (3) for the sub-periods defined by the breaks in (3).

Table 2.2: Estimated Rates of Growth and Volatility for the U.K.

			GMM Models		
				Test for Additional Breaks	
	Random Walk	No Break	Volatility Break	Period 1	Period 2
	(1)	(2)	(3)	(4)	(5)
Constant	0.6012	0.6236	0.6436	0.6129	0.1833
(P-values in Parentheses)	(0.067)	(0.000)	(0.000)	(0.000)	(0.117)
AR(1)		-0.0312	-0.0265	-0.1034	0.7027
		(0.712)	(0.754)	(0.189)	(0.000)
Std. Deviation (Pre-Break)	0.9730	0.8393	1.0565	1.1981	0.2800
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)			0.3779	0.7784	0.3785
			(0.000)	(0.000)	(0.000)
Wald Test Values					
Sup			42.4025	9.4951	2.4822
			(0.000)	(0.106)	(0.866)
Ave			18.0907	2.5466	0.5272
			(0.000)	(0.099)	(0.774)
Exp			19.5112	2.4278	0.9473
			(0.000)	(0.278)	(0.753)
J-Statistic			1.9908	0.3729	4.5696
Break Dates					
Wald Statistic			1992Q2	1981Q2	2000Q4

Values are calculated for these parameters when allowing for breaks in the specific parameter. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the mean, represents the mean before the break. The second reported value for each parameter when allowing for breaks estimates the value after the break.
*Wald tests here signify a test of the restriction rather than a test for the break date. The parameter values are those from the estimated break date supported by the model.
Column (1) represents a random walk model of growth, and estimates the standard deviation of the residual values.
Column (2) estimates a basic ARIMA(1,1,0) model of output using GMM Equations (1) and (2) with no break.
Column (3) estimates an ARIMA(1,1,0) model of output with parameter breaks std. deviation.
Columns (4) and (5) estimate the same model as column (3) for the sub-periods defined by the breaks in (3).

Table 2.3: Estimated Rates of Growth and Volatility for Australia

			GMM Models		
				Test for Additional Breaks	
	Random Walk	No Break	Volatility Break	Period 1	Period 2
	(1)	(2)	(3)	(4)	(5)
Constant	0.8849	0.9036	0.8600	0.9988	0.4442
(P-values in Parentheses)	(0.081)	(0.000)	(0.000)	(0.000)	(0.000)
AR(1)		-0.0315	-0.0172	-0.1022	0.3949
		(0.746)	(0.860)	(0.331)	(0.000)
Std. Deviation (Pre-Break)	1.2814	0.9954	1.3994	1.5279	0.6978
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)			0.6111	1.2342	0.2974
			(0.000)	(0.000)	(0.000)
Wald Test Values					
Sup			67.4453	4.6710	50.1581
			(0.000)	(0.542)	(0.000)
Ave			29.4845	0.9859	21.6972
			(0.000)	(0.507)	(0.000)
Exp			32.1355	1.6130	13.0583
			(0.001)	(0.502)	(0.000)
J-Statistic			2.6291	2.4596	4.2680
Break Dates					
Wald Statistic			1985Q2	1975Q4	2004Q1

Values are calculated for these parameters when allowing for breaks in the specific parameter. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the mean, represents the mean before the break. The second reported value for each parameter when allowing for breaks estimates the value after the break.
*Wald tests here signify a test of the restriction rather than a test for the break date. The parameter values are those from the estimated break date supported by the model.
Column (1) represents a random walk model of growth, and estimates the standard deviation of the residual values.
Column (2) estimates a basic ARIMA(1,1,0) model of output using GMM Equations (1) and (2) with no break.
Column (3) estimates an ARIMA(1,1,0) model of output with parameter breaks std. deviation.
Columns (4) and (5) estimate the same model as column (3) for the sub-periods defined by the breaks in (3).

Table 2.4: Markov-Switching Estimation Results

	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
U.S.												
μ_{11}	1.134	*(0.111)*	0.903	*(0.098)*	1.130	*(0.105)*	0.724	*(0.081)*	1.374	*(0.136)*	1.293	*(0.182)*
μ_{21}	-0.033	*(0.250)*	-1.233	*(0.472)*	-0.012	*(0.240)*	0.861	*(0.139)*	-0.148	*(0.210)*	-0.189	*(0.282)*
μ_{12}									0.903	*(0.064)*	0.877	*(0.078)*
μ_{22}									0.192	*(0.136)*	0.168	*(0.190)*
ϕ_1			0.401	*(0.069)*			0.328	*(0.065)*			0.142	*(0.102)*
σ_1^2	0.685	*(0.074)*	0.668	*(0.080)*	0.652	*(0.080)*	0.255	*(0.039)*	0.893	*(0.133)*	0.930	*(0.157)*
σ_2^2					0.806	*(0.192)*	1.249	*(0.148)*	0.168	*(0.031)*	0.182	*(0.034)*
p_{11}	0.908	*(0.040)*	0.966	*(0.022)*	0.910	*(0.037)*	0.993	*(0.009)*	0.902	*(0.034)*	0.913	*(0.037)*
p_{22}	0.770	*(0.091)*	0.232	*(0.218)*	0.778	*(0.091)*	0.995	*(0.006)*	0.782	*(0.071)*	0.758	*(0.095)*
q_{11}									0.994	*(0.006)*	0.995	*(0.006)*
q_{22}									0.993	*(0.010)*	0.994	*(0.008)*
Log L	-109.578		-99.233		-109.215		-76.735		-76.034		-74.374	
U.K.												
μ_{11}	0.733	*(0.082)*	0.772	*(0.060)*	0.732	*(0.037)*	0.731	*(0.037)*	0.762	*(0.184)*	0.544	*(0.189)*
μ_{21}	-0.657	*(0.470)*	-0.532	*(0.265)*	0.547	*(0.094)*	0.550	*(0.097)*	-0.057	*(0.616)*	0.643	*(0.561)*
μ_{12}									0.854	*(0.081)*	0.785	*(0.043)*
μ_{22}									0.615	*(0.057)*	-0.423	*(0.138)*
ϕ_1			-0.215	*(0.074)*			0.025	*(0.070)*			0.013	*(0.071)*
σ_1^2	0.803	*(0.105)*	0.737	*(0.087)*	0.079	*(0.015)*	0.078	*(0.015)*	1.195	*(0.166)*	1.644	*(0.216)*
σ_2^2					1.329	*(0.155)*	1.343	*(0.159)*	0.065	*(0.014)*	0.170	*(0.023)*
p_{11}	0.968	*(0.023)*	0.961	*(0.020)*	0.976	*(0.021)*	0.976	*(0.022)*	0.955	*(0.066)*	0.973	*(0.016)*
p_{22}	0.710	*(0.208)*	0.756	*(0.133)*	0.993	*(0.008)*	0.993	*(0.008)*	0.917	*(0.071)*	0.880	*(0.110)*
q_{11}									0.994	*(0.007)*	0.998	*(0.003)*
q_{22}									0.979	*(0.018)*	0.997	*(0.004)*
Log L	-101.215		-97.260		-58.549		-58.253		-56.260		-56.754	
Australia												
μ_{11}	1.004	*(0.123)*	1.064	*(0.094)*	0.938	*(0.144)*	0.932	*(0.161)*	0.949	*(0.211)*	0.954	*(0.261)*
μ_{21}	-0.413	*(0.552)*	-0.354	*(0.304)*	0.828	*(0.068)*	0.829	*(0.072)*	0.865	*(1.224)*	0.807	*(1.711)*
μ_{12}									0.949	*(0.067)*	0.950	*(0.064)*
μ_{22}									-0.161	*(0.241)*	-0.174	*(0.235)*
ϕ_1			-0.176	*(0.082)*			0.081	*(0.075)*			-0.040	*(0.077)*
σ_1^2	1.120	*(0.165)*	1.023	*(0.134)*	2.077	*(0.299)*	2.156	*(0.319)*	2.088	*(0.297)*	2.089	*(0.303)*
σ_2^2					0.416	*(0.062)*	0.403	*(0.060)*	0.293	*(0.050)*	0.290	*(0.050)*
p_{11}	0.962	*(0.043)*	0.949	*(0.037)*	0.994	*(0.007)*	0.994	*(0.007)*	0.970	*(0.025)*	0.969	*(0.026)*
p_{22}	0.584	*(0.257)*	0.646	*(0.149)*	0.994	*(0.007)*	0.994	*(0.007)*	0.769	*(0.174)*	0.764	*(0.180)*
q_{11}									0.994	*(0.007)*	0.994	*(0.007)*
q_{22}									0.994	*(0.007)*	0.994	*(0.007)*
Log L	-121.068		-118.867		-97.724		-96.745		-91.848		-91.319	

Standard Errors are in parentheses
CV-NoAR: No switching variance, no AR Coefficient (Basic MS Model), CV-AR(1): No switching variance, AR Coefficient
TS-NoAR: Tied Switching Variance, no AR Coefficient, TS-AR(1): Tied Switching Variance, AR Coefficient
IS-NoAR: Independent Switching Variance, no AR Coefficient, IS-AR(1): Independent Switching Variance, AR coefficient

Table 2.5: Simulation Model Specifications

	Linear Simulations
Model 1	A random walk model.
Model 2	GMM estimated model without breaks in volatility.
Model 3	Apply pre-break standard deviation to the entire simulation.
Model 4	Apply post-break standard deviation to the entire simulation.
Model 5	Apply pre- and post-break standard deviations to respective sample lengths(“Switching”).
Model 6	Apply pre- and post-break standard deviations to opposite sample lengths (“Reversing”).
	Markov-Switching Simulations
Model 7	MS-CV-NoAR model with constant volatility and no linear structure.
Model 8	MS-CV-AR(1): Model 7 with linear structure.
Model 9	MS-TS-NoAR model with tied switching and no linear structure.
Model 10	MS-TS-AR(1): Model 9 with linear structure.
Model 11	MS-IS-NoAR model with independent switching and no linear structure.
Model 12	MS-IS-AR(1): Model 11 with linear structure.

Table 2.6: Generalized Q-Test Statistics: U.S. Simulation Features vs. Actual Data Features

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.012	0.008	**0.265**	**0.751**	0.010	**0.442**	**0.118**	0.027	**0.164**
GMM No Breaks (2)	0.006	0.004	**0.231**	**0.960**	**0.068**	**0.946**	**0.089**	0.039	**0.271**
Period 1 Volatility	0.041	**0.653**	0.007	**0.134**	**0.398**	**0.242**	0.009	**0.228**	0.022
Period 2 Volatility	0.000	0.000	**0.393**	**0.361**	0.000	0.000	**0.572**	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.506**	**0.733**	**0.288**	**0.718**	**0.884**	**0.782**	**0.134**	**0.851**	**0.339**
SD Reversing (Vol 2/Vol 1)	**0.580**	**0.672**	**0.367**	**0.825**	**0.907**	**0.830**	**0.164**	**0.806**	**0.431**
MS Constant Vol-NoAR	0.000	**0.061**	0.000	0.024	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.415**	0.000	**0.123**	0.013	0.036	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.142**	0.000	0.027	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	**0.204**	0.000	**0.402**	**0.245**	**0.404**	0.001	0.000	0.007
MS Indep Switching-NoAR	**0.841**	**0.625**	**0.856**	**0.700**	**0.850**	**0.823**	**0.583**	**0.724**	**0.815**
MS Indep Switching-AR(1)	**0.930**	**0.799**	**0.867**	**0.695**	**0.825**	**0.773**	**0.666**	**0.788**	**0.862**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 2.7: Generalized Q-Test Statistics: U.K. Simulation Features vs. Actual Data Features

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.001	0.027	0.002	0.000	0.000	**0.131**	0.001	0.000
GMM No Breaks (2)	0.000	0.000	**0.209**	0.006	0.000	0.000	**0.201**	0.000	0.000
Period 1 Volatility	0.000	0.004	0.010	0.001	0.001	0.000	**0.147**	0.001	0.000
Period 2 Volatility	0.000	0.000	**0.414**	0.000	0.000	0.000	**0.170**	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.028	0.037	**0.133**	0.020	0.008	0.003	**0.302**	0.021	0.008
SD Reversing (Vol 2/Vol 1)	**0.507**	**0.370**	**0.805**	**0.322**	**0.123**	**0.085**	**0.713**	**0.335**	**0.250**
MS Constant Vol-NoAR	0.000	**0.595**	0.000	0.000	0.000	0.000	0.003	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.134**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.735**	0.000	**0.626**	0.017	**0.447**	**0.301**	0.000	**0.436**
MS Tied Switching-AR(1)	0.000	**0.806**	0.000	**0.610**	0.030	**0.539**	**0.340**	0.000	**0.448**
MS Indep Switching-NoAR	**0.162**	**0.485**	**0.152**	**0.326**	**0.128**	**0.122**	**0.607**	**0.352**	**0.307**
MS Indep Switching-AR(1)	0.000	**0.781**	0.000	**0.058**	**0.103**	**0.757**	**0.121**	**0.065**	**0.301**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 2.8: Generalized Q-Test Statistics: Australia Simulation Features vs. Actual Data Features

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	**0.515**	**0.407**	**0.534**	**0.159**	**0.778**	**0.459**	**0.355**	**0.294**	**0.358**
GMM No Breaks (2)	**0.104**	0.046	**0.500**	**0.055**	**0.276**	**0.060**	**0.448**	**0.091**	**0.053**
Period 1 Volatility	**0.405**	**0.328**	**0.457**	**0.599**	**0.597**	**0.970**	**0.225**	**0.320**	**0.525**
Period 2 Volatility	0.000	0.000	**0.443**	0.000	0.000	0.000	**0.411**	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.842**	**0.668**	**0.791**	**0.656**	**0.862**	**0.941**	**0.584**	**0.597**	**0.848**
SD Reversing (Vol 2/Vol 1)	**0.849**	**0.722**	**0.782**	**0.633**	**0.895**	**0.926**	**0.452**	**0.624**	**0.802**
MS Constant Vol-NoAR	0.000	**0.075**	0.000	**0.777**	**0.122**	**0.811**	**0.052**	0.000	**0.147**
MS Constant Vol-AR(1)	0.000	0.000	**0.242**	0.018	0.000	0.000	**0.274**	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.082**	0.000	**0.766**	**0.128**	**0.779**	**0.066**	0.000	**0.166**
MS Tied Switching-AR(1)	0.000	0.000	**0.349**	0.009	0.000	0.000	**0.060**	0.000	0.000
MS Indep Switching-NoAR	**0.976**	**0.911**	**0.894**	**0.620**	**0.691**	**0.729**	**0.527**	**0.905**	**0.820**
MS Indep Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.576**	0.000	0.000

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 2.9: Linear Model Simulation Peak to Trough and Trough to Peak Characteristics

		Actual	RW		No Breaks		High Volatility		Low Volatility		Switching Volatility		Reversing Volatility	
U.S.														
P→T	Durations	2.889	2.647	**76%**	2.989	**48%**	3.475	**13%**	2.393	**82%**	3.375	**24%**	3.264	**35%**
	Amplitudes	-2.238	-1.323	0%	-1.428	2%	-2.334	**55%**	-0.582	0%	-2.167	**40%**	-1.963	**29%**
	Cumulation	-2.938	-1.933	**10%**	-2.557	**28%**	-5.101	**89%**	-0.755	1%	-4.723	**76%**	-4.238	**62%**
	Excess	0.024	0.001	**64%**	0.000	**65%**	-0.001	**63%**	-0.001	**70%**	-0.002	**61%**	0.000	**61%**
T→P	Durations	21.700	27.416	**34%**	23.625	**50%**	14.922	**94%**	65.568	8%	18.472	**75%**	26.255	**41%**
	Amplitudes	22.135	25.647	**43%**	22.517	**59%**	16.184	**91%**	51.118	**14%**	18.723	**77%**	24.318	**50%**
	Cumulation	402.808	636.306	**43%**	489.422	**58%**	208.897	**93%**	2741.296	**15%**	340.732	**73%**	713.492	**43%**
	Excess	1.116	0.024	98%	0.020	98%	0.008	100%	0.009	**87%**	0.110	97%	-0.186	98%
U.K.														
P→T	Durations	4.167	3.053	97%	2.780	99%	3.053	97%	2.085	100%	3.032	**95%**	2.969	**91%**
	Amplitudes	-2.685	-1.609	0%	-1.201	0%	-1.706	1%	-0.312	0%	-1.682	2%	-1.592	**5%**
	Cumulation	-7.199	-2.837	1%	-1.865	0%	-2.974	1%	-0.329	0%	-2.954	1%	-2.751	3%
	Excess	-0.061	-0.002	**19%**	-0.001	**16%**	0.001	**20%**	0.001	**8%**	-0.002	**25%**	0.005	**29%**
T→P	Durations	31.000	16.875	98%	22.213	**88%**	16.691	98%	73.224	**22%**	17.200	96%	22.756	**78%**
	Amplitudes	24.250	13.590	98%	16.313	**92%**	14.082	97%	46.043	**31%**	14.397	**95%**	17.448	**80%**
	Cumulation	522.574	194.506	97%	320.927	**87%**	199.345	97%	2476.760	**31%**	213.969	**94%**	485.519	**78%**
	Excess	-0.562	0.010	3%	0.017	5%	0.016	3%	-0.024	**22%**	0.031	**6%**	-0.094	**18%**
Australia														
P→T	Durations	3.571	2.708	**95%**	2.540	97%	3.038	**85%**	2.178	98%	2.956	**84%**	2.934	**85%**
	Amplitudes	-1.791	-1.588	**27%**	-1.264	**7%**	-2.260	**85%**	-0.576	0%	-2.106	**68%**	-2.047	**64%**
	Cumulation	-3.667	-2.395	**11%**	-1.737	3%	-3.947	**48%**	-0.650	0%	-3.658	**38%**	-3.520	**36%**
	Excess	0.074	-0.001	**80%**	0.000	**82%**	0.000	**77%**	0.000	**86%**	-0.002	**73%**	0.002	**72%**
T→P	Durations	16.500	24.681	**18%**	31.959	**9%**	16.889	**56%**	57.839	**11%**	20.403	**45%**	21.309	**42%**
	Amplitudes	20.044	25.792	**33%**	31.391	**22%**	19.090	**66%**	49.934	**19%**	21.857	**54%**	22.553	**51%**
	Cumulation	210.276	562.230	**19%**	884.700	**12%**	274.439	**50%**	2211.311	**16%**	430.017	**44%**	445.397	**42%**
	Excess	0.798	0.009	**92%**	0.048	**87%**	0.026	**95%**	0.025	**80%**	0.101	**87%**	-0.086	**91%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) "RW" is a Random Walk model with constant mean and variance terms.
Column (2) "No Breaks" is an ARIMA(1,1,0) model of output with a constant mean, variance, and AR(1) term.
Columns (3) - (6) are all simulated using estimated parameters from column (3) of Tables 2.1 - 2.3, applying the volatility measures to different parts of the simulations.
Column (3) "High Volatility" is an ARIMA(1,1,0) simulation of output, using the first period variance over the entire time period.
Column (4) "Low Volatility" is an ARIMA(1,1,0) simulation of output, using the second period variance over the entire time period.
Column (5) "Switching Volatility" is an ARIMA(1,1,0) simulation of output, switching the variances (σ_1^2 to σ_2^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.
Column (6) "Reversing Volatility" is an ARIMA(1,1,0) simulation of output, reversing the variances (σ_2^2 to σ_1^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.

Table 2.10: Markov-Switching Simulation Peak to Trough and Trough to Peak Characteristics

		Actual	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
U.S.														
P→T	Durations	2.889	3.920	3%	3.368	**18%**	3.979	3%	3.700	**7%**	3.396	**33%**	3.387	**34%**
	Amplitudes	-2.238	-1.987	**23%**	-2.070	**31%**	-2.106	**34%**	-2.155	**39%**	-1.836	**35%**	-2.006	**43%**
	Cumulation	-2.938	-4.858	**88%**	-4.290	**79%**	-5.193	**91%**	-5.007	**91%**	-4.329	**58%**	-4.792	**62%**
	Excess	0.024	0.001	**63%**	-0.008	**68%**	0.001	**62%**	-0.104	**91%**	0.002	**61%**	0.004	**60%**
T→P	Durations	21.700	13.708	97%	14.628	95%	13.837	97%	18.882	**76%**	34.940	**36%**	36.529	**34%**
	Amplitudes	22.135	9.230	100%	12.466	99%	9.467	100%	14.249	**94%**	32.457	**33%**	38.176	**22%**
	Cumulation	402.808	106.108	99%	155.133	97%	109.801	99%	238.215	**88%**	1152.206	**35%**	1403.039	**29%**
	Excess	1.116	0.003	100%	0.006	100%	-0.001	100%	-0.058	100%	0.039	**91%**	-0.014	**90%**
U.K.														
P→T	Durations	4.167	4.156	**54%**	4.252	**50%**	4.573	**36%**	4.559	**37%**	3.295	**89%**	5.346	**65%**
	Amplitudes	-2.685	-2.245	**14%**	-2.055	4%	-2.378	**22%**	-2.412	**25%**	-1.830	5%	-2.333	**21%**
	Cumulation	-7.199	-5.951	**24%**	-5.409	**16%**	-6.208	**28%**	-6.356	**29%**	-3.544	4%	-9.990	**31%**
	Excess	-0.061	0.000	**21%**	0.001	**19%**	-0.009	**33%**	-0.028	**39%**	0.002	**24%**	0.000	**23%**
T→P	Durations	31.000	10.630	100%	10.546	100%	22.383	**86%**	21.992	**88%**	17.095	**93%**	12.113	99%
	Amplitudes	24.250	6.777	100%	5.635	100%	10.871	98%	11.195	98%	15.195	**91%**	10.718	98%
	Cumulation	522.574	55.755	100%	44.641	100%	219.544	**94%**	225.886	**93%**	330.193	**89%**	145.784	98%
	Excess	-0.562	0.000	0%	-0.006	0%	-0.021	5%	-0.043	**6%**	0.017	**14%**	0.007	3%
Australia														
P→T	Durations	3.571	3.713	**50%**	2.383	97%	3.701	**49%**	2.338	98%	4.263	**33%**	6.005	0%
	Amplitudes	-1.791	-1.174	4%	-0.672	1%	-1.170	4%	-0.677	1%	-1.983	**57%**	-7.510	100%
	Cumulation	-3.667	-2.681	**19%**	-0.877	1%	-2.658	**19%**	-0.934	1%	-5.951	**63%**	-34.229	100%
	Excess	0.074	0.000	**89%**	0.000	**92%**	0.001	**89%**	-0.042	97%	0.000	**82%**	0.006	**72%**
T→P	Durations	16.500	26.525	**18%**	46.363	**8%**	26.792	**17%**	49.289	**9%**	29.568	**13%**	8.676	100%
	Amplitudes	20.044	14.770	**84%**	37.838	**22%**	14.867	**84%**	40.135	**21%**	36.304	**14%**	27.155	5%
	Cumulation	210.276	367.067	**45%**	1435.410	**13%**	369.033	**44%**	1584.259	**14%**	955.982	**10%**	169.644	**86%**
	Excess	0.798	-0.003	98%	0.000	**83%**	0.000	98%	-0.017	**83%**	-0.279	**88%**	0.223	**84%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) CV-NoAR is a Markov-switching model with a constant variance and no AR(1) term.
Column (2) CV-AR(1) is a Markov-switching model with a constant variance and AR(1) term.
Column (3) TS-NoAR is a Markov-switching model with a tied switching variance and no AR(1) term.
Column (4) TS-AR(1) is a Markov-switching model with a tied switching variance and AR(1) term.
Column (5) IS-NoAR is a Markov-switching model with an independent mean and variance and no AR(1) term.
Column (6) IS-AR(1) is a Markov-switching model with an independent mean and variance with an AR(1) term.

Table 2.11: Estimated Values for Regressions Including Recession Dummy Variables

	Recession Dummy Variable			
U.S.	NBER	IS-NoAR	BBQ	None
AR(1) Coefficient	0.101	0.0792	0.1548	0.3367
S.E.	0.0615	0.0547	0.0552	0.0606
Box-Ljung (p-value)	0.3252	0.7586	0.4438	0.3195
Durbin-Watson	1.9489	1.904	1.8789	2.0458
U.K.	IS-NoAR[a]	CV-AR(1)[b]	BBQ	None
AR(1) Coefficient	-0.2227	-0.1431	-0.2266	-0.0548
S.E.	0.0605	0.068	0.0615	0.0693
Box-Ljung (p-value)	0.7444	0.5593	0.5609	0.152
Durbin-Watson	1.9477	1.9508	1.9936	1.9508
Australia	CV-AR(1)[a]	IS-NoAR[a]	BBQ	None
AR(1) Coefficient	-0.1402	-0.1958	-0.191	-0.0435
S.E.	0.0688	0.068	0.065	0.072
Box-Ljung (p-value)	0.226	0.1235	0.1056	0.0655
Durbin-Watson	2.0726	2.0409	2.043	1.996

[a] Probability of being in low-mean state.
[b] Probability of being in low-mean state.

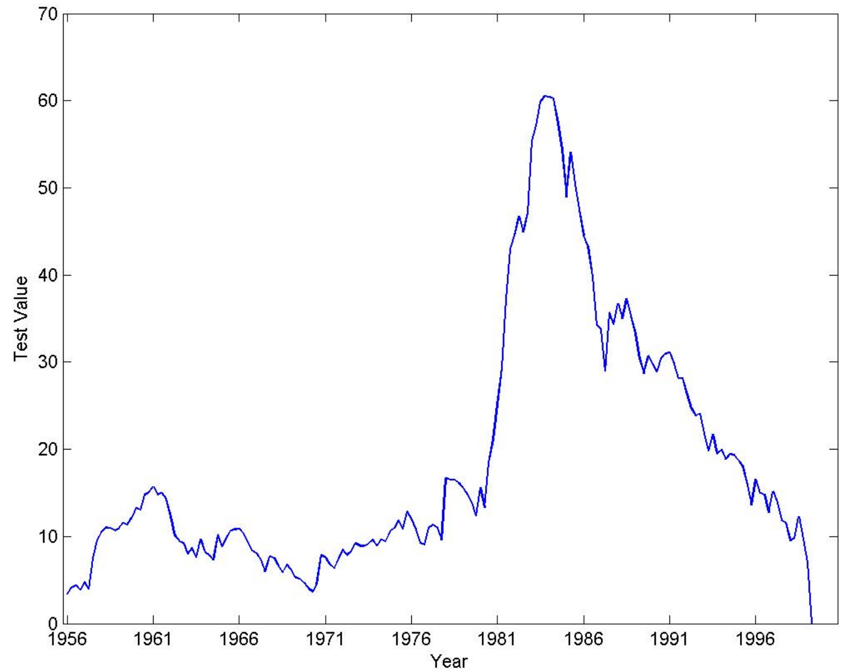


Figure 2.1: U.S. Wald Sup Tests for a Single Break

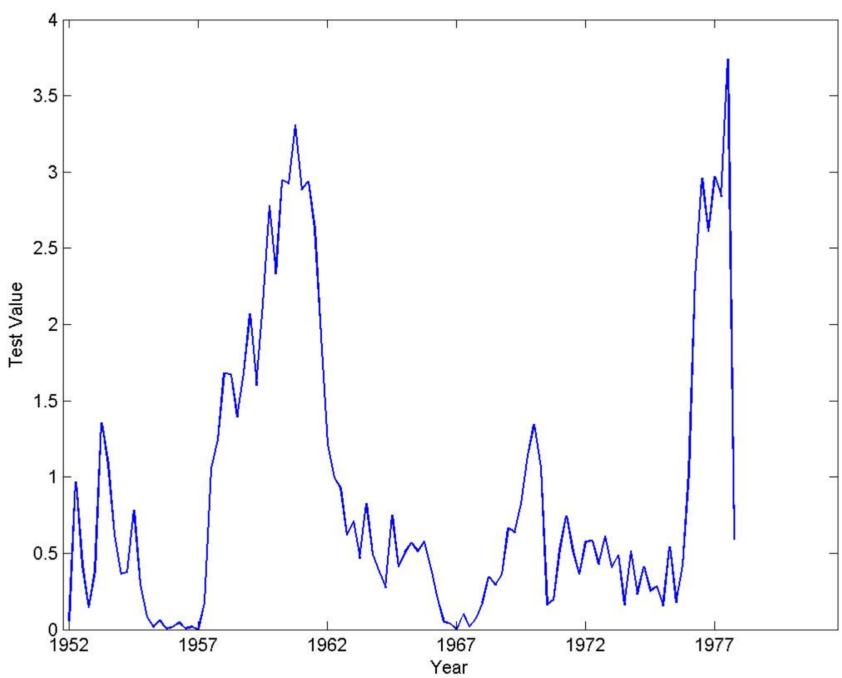


Figure 2.2: U.S. Wald Sup Tests for Period Before Break

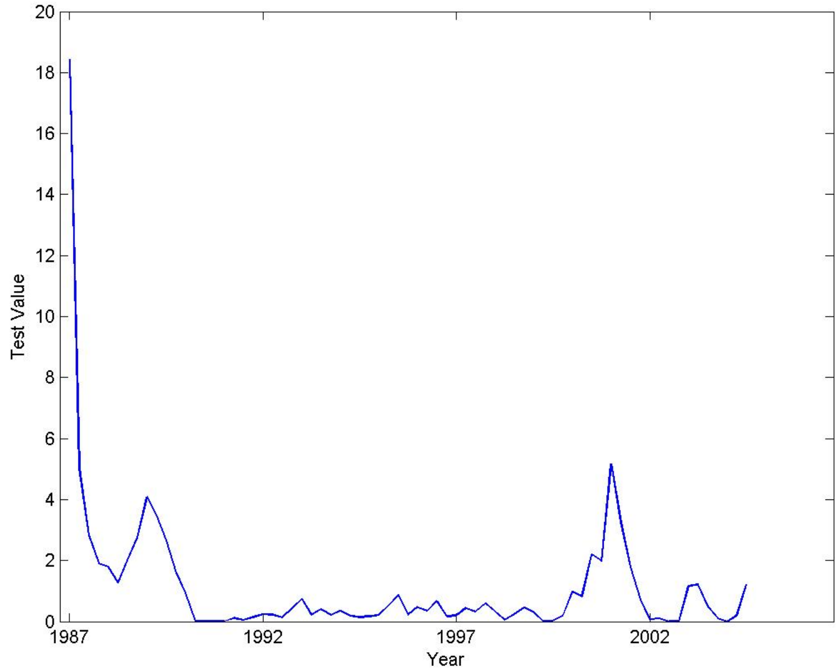


Figure 2.3: U.S. Wald Sup Tests for Period After Break

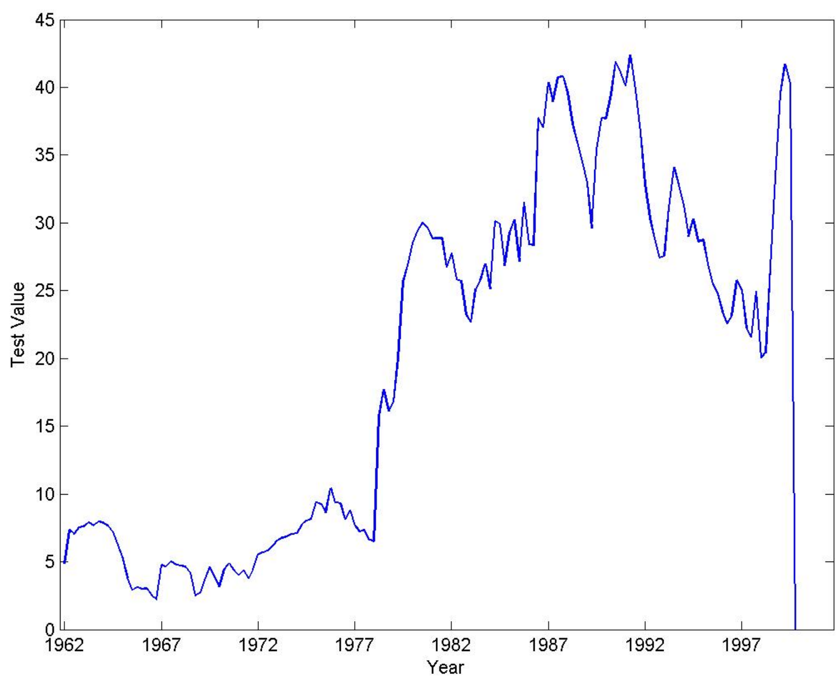


Figure 2.4: U.K. Wald Sup Tests for a Single Break

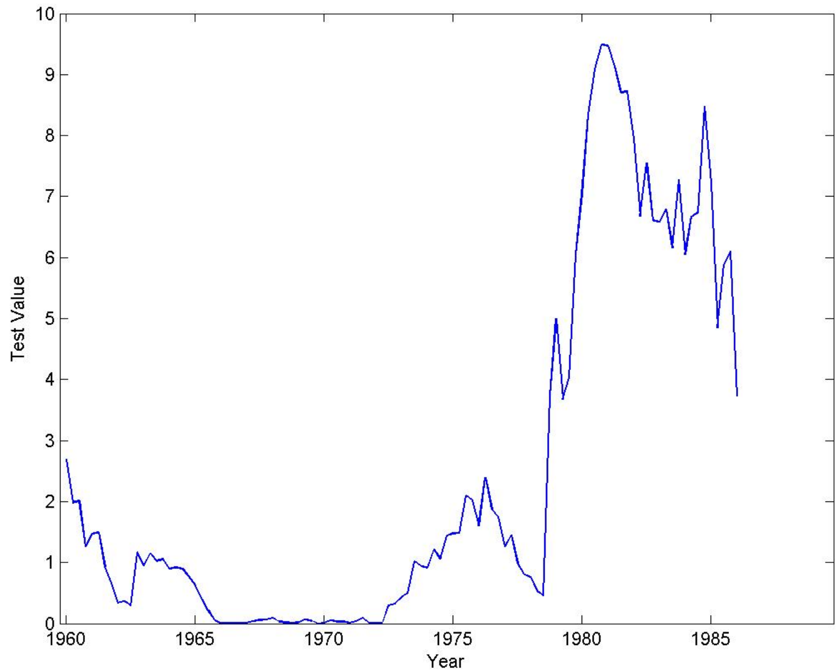


Figure 2.5: U.K. Wald Sup Tests for Period Before Break

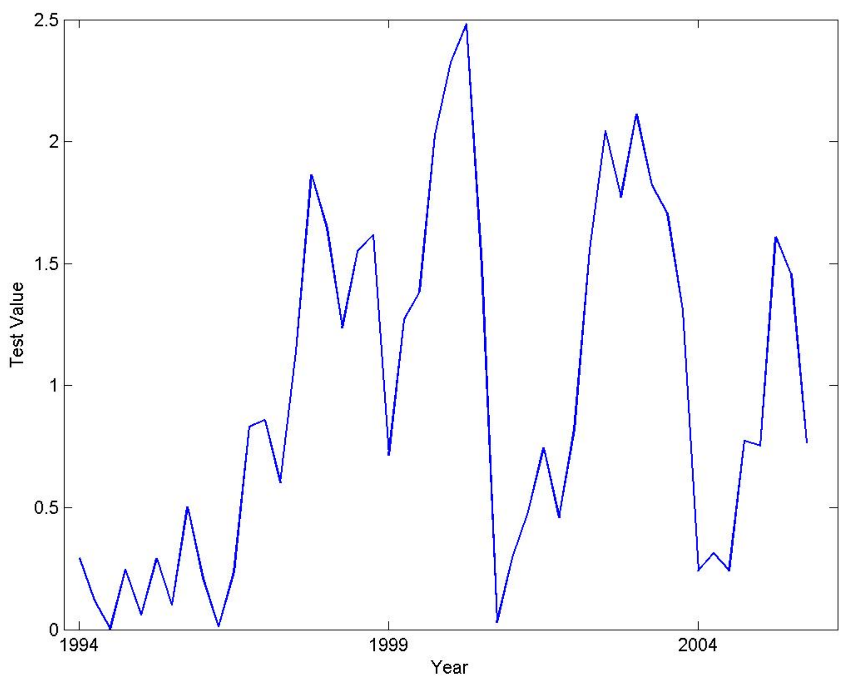


Figure 2.6: U.K. Wald Sup Tests for Period After Break

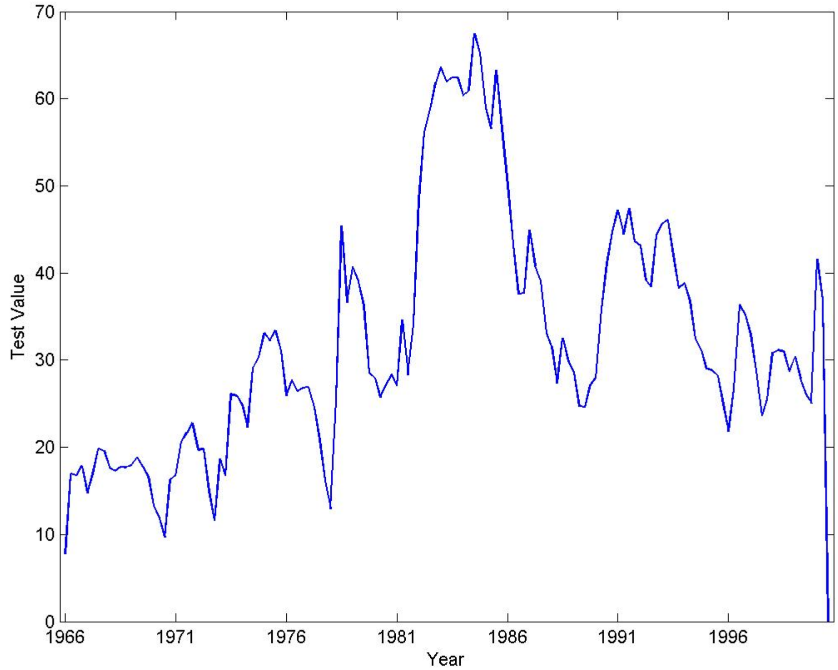


Figure 2.7: Australia Wald Sup Tests for a Single Break

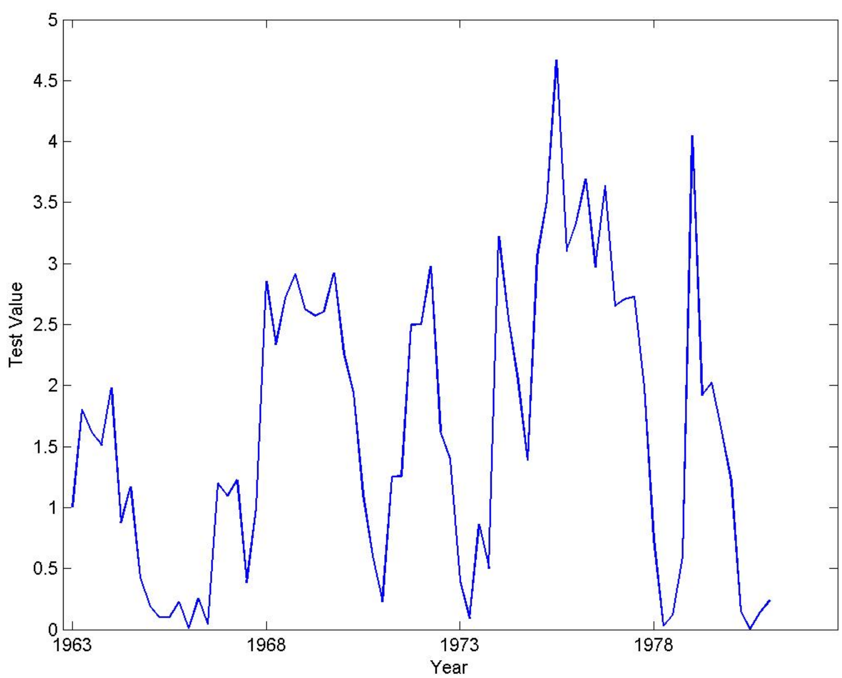


Figure 2.8: Australia Wald Sup Tests for Period Before Break

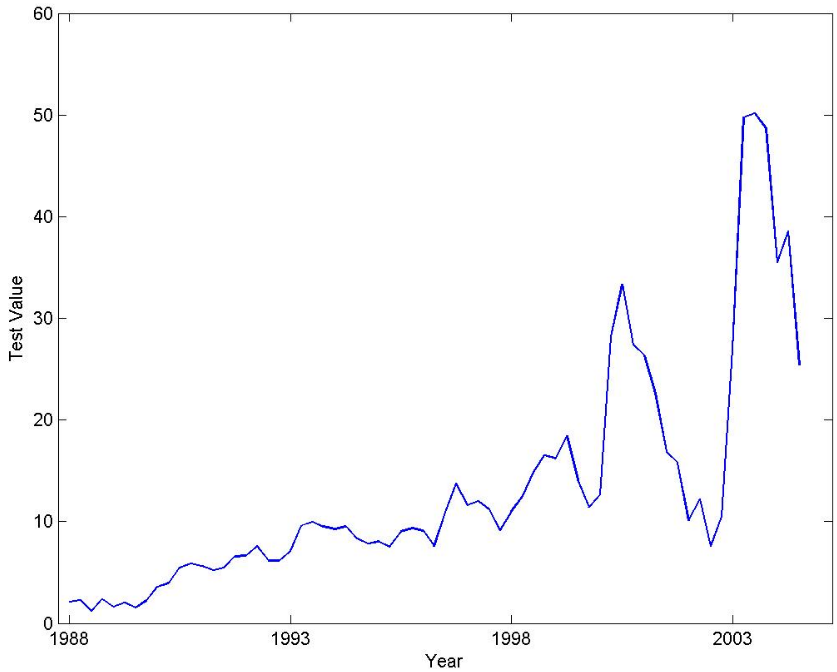


Figure 2.9: Australia Wald Sup Tests for Period After Break

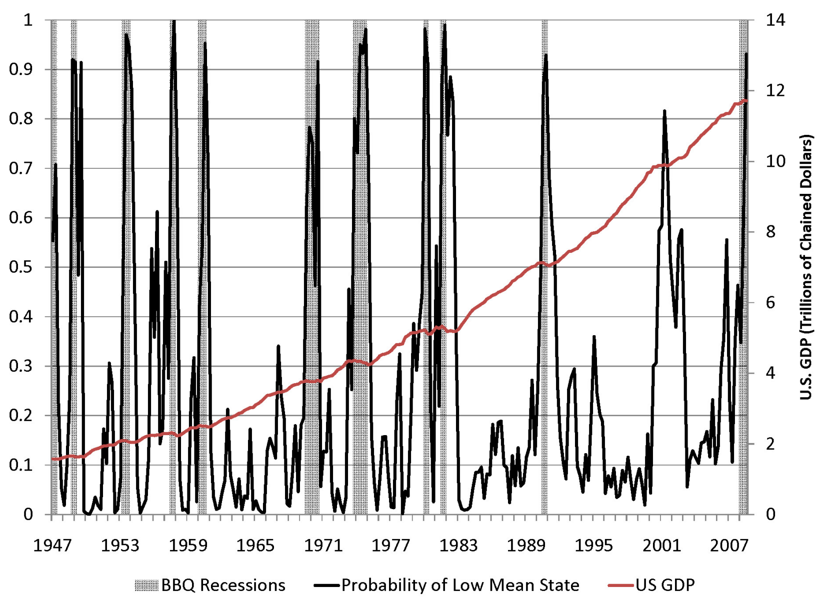


Figure 2.10: U.S. Markov-Switching Model with Constant Variance and No AR(1) Term: Probabilities of Being in State 2 (Low-Mean State)

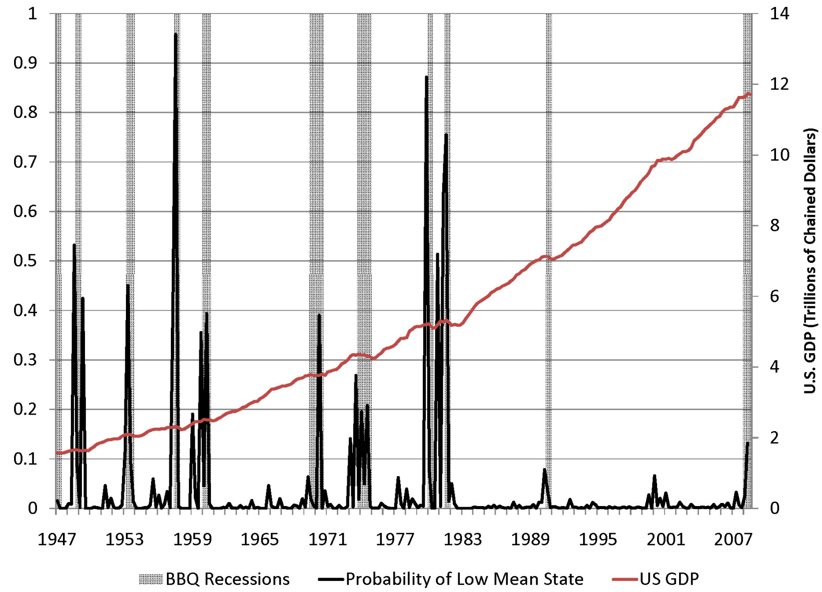


Figure 2.11: U.S. Markov-Switching Model with Constant Variance and AR(1) Term: Probabilities of Being in State 2 (Low-Mean State)

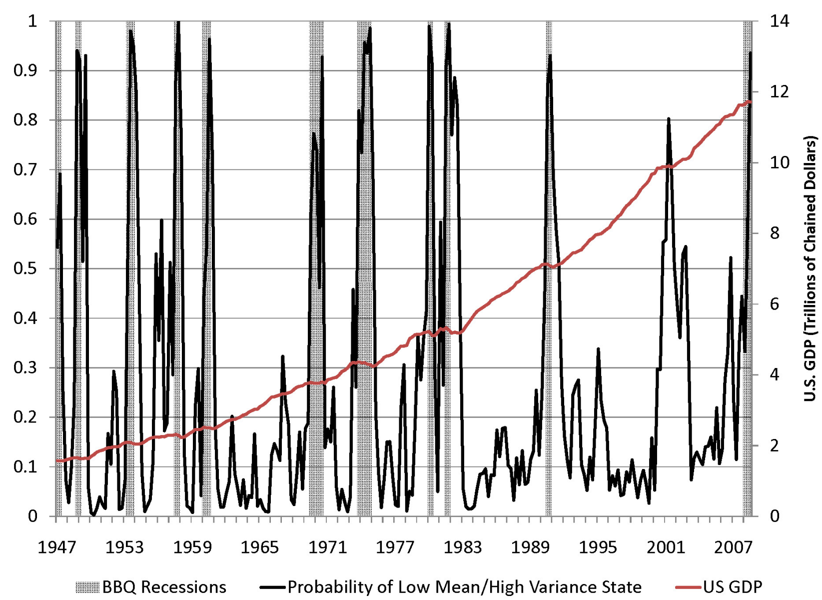


Figure 2.12: U.S. Markov-Switching Model with Tied Switching Variance and No AR(1) Term: Probabilities of Being in State 2 (Low-Mean/High-Variance State)

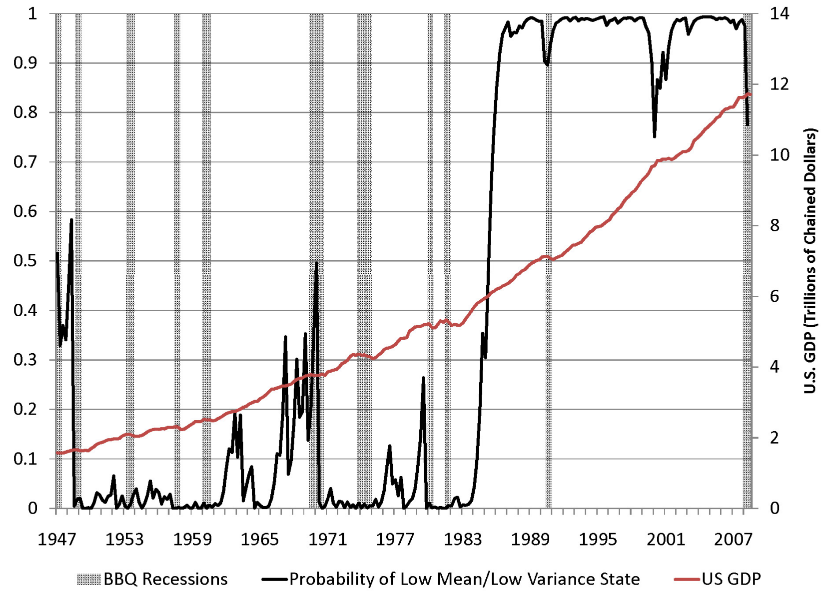


Figure 2.13: U.S. Markov-Switching Model with Tied Switching Variance and AR(1) Term: Probabilities of Being in State 2 (Low-Mean/Low-Variance State)

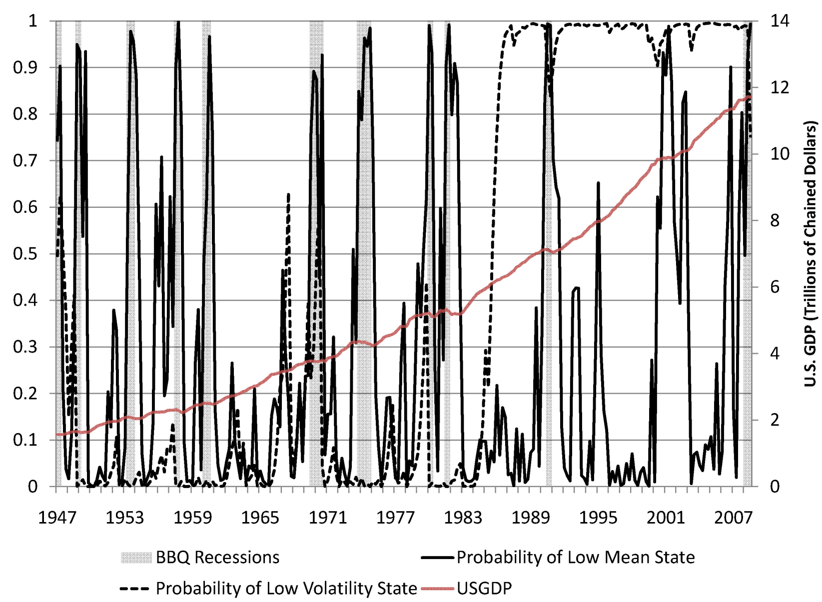


Figure 2.14: U.S. Markov-Switching Model with Independent Switching Variance and No AR(1) Term: Probabilities of Being in Low-Mean State and Low-Volatility State

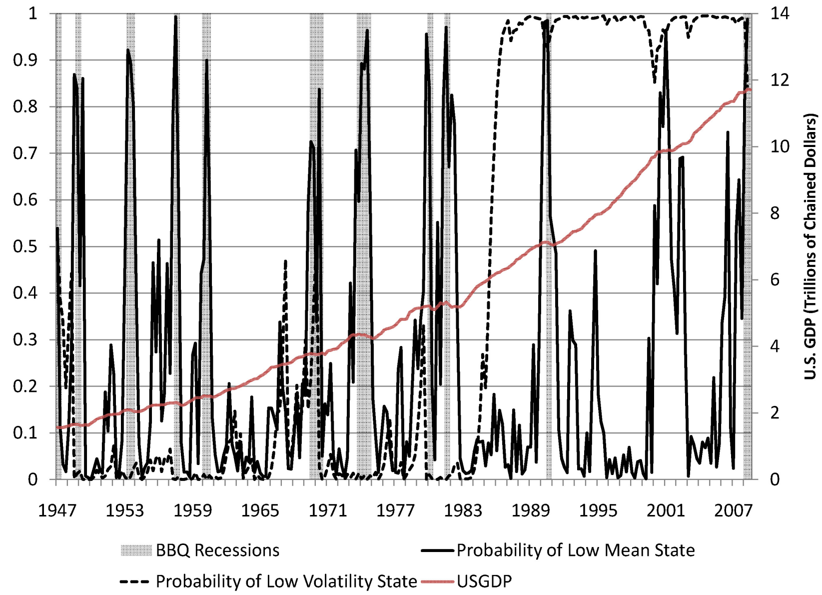


Figure 2.15: U.S. Markov-Switching Model with Independent Switching Variance and AR(1) Term: Probabilities of Being in Low-Mean State and Low-Volatility State

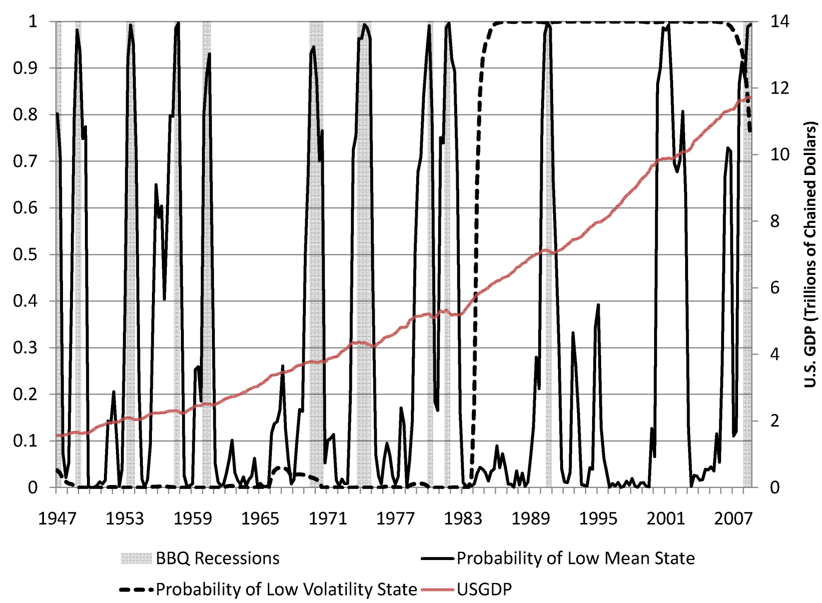


Figure 2.16: U.S. Markov-Switching Model with Independent Switching Variance and No AR(1) Term: Smoothed Probabilities of Being in Low-Mean State and Low-Volatility State

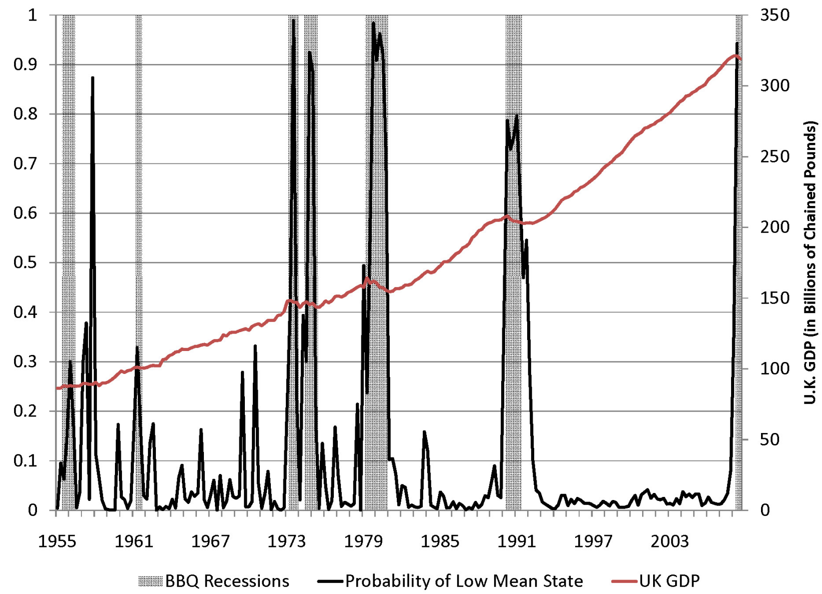


Figure 2.17: U.K. Markov-Switching Model with Constant Variance and AR(1) Term: Probabilities of Being in State 2 (Low-Mean)

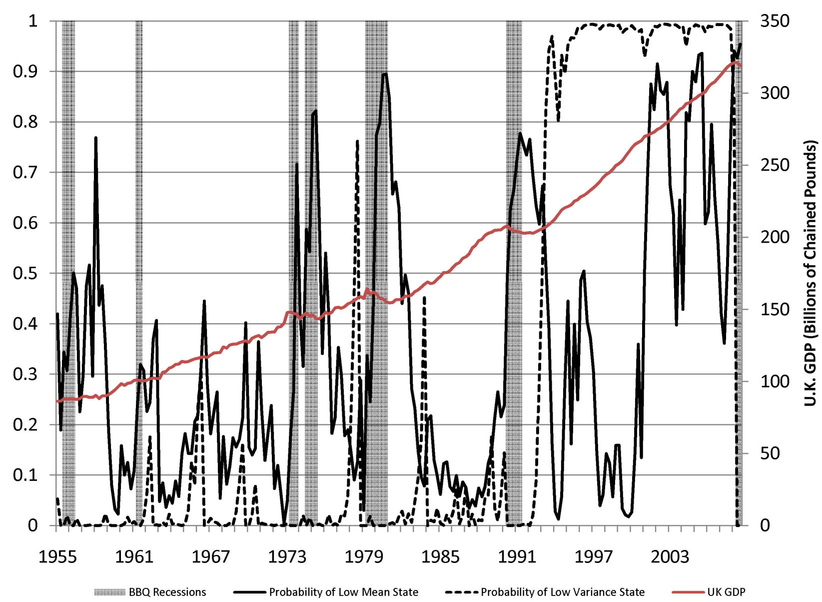


Figure 2.18: U.K. Markov-Switching Model with Independent Switching Variance and No AR(1) Term: Probabilities of Being in Low-Mean State and Low-Variance State

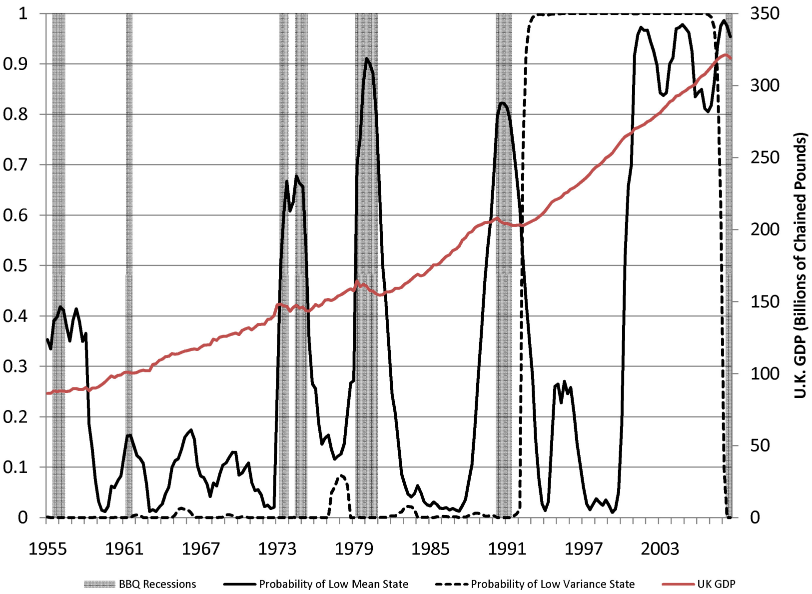


Figure 2.19: U.K. Markov-Switching Model with Independent Switching Variance and No AR(1) Term: Smoothed Probabilities of Being in Low-Mean State and Low-Variance State

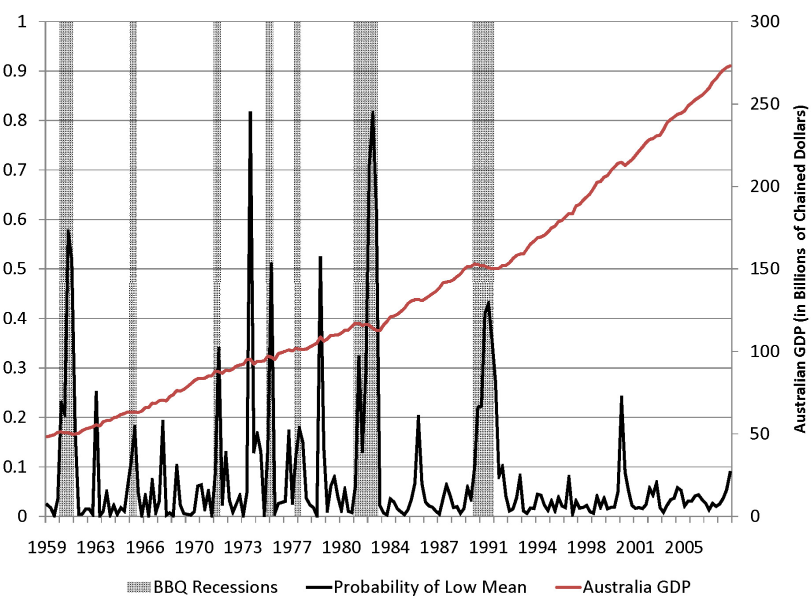


Figure 2.20: Australia Markov-Switching Model with Constant Variance and No AR(1) Term: Probabilities of Being in Low-Mean State

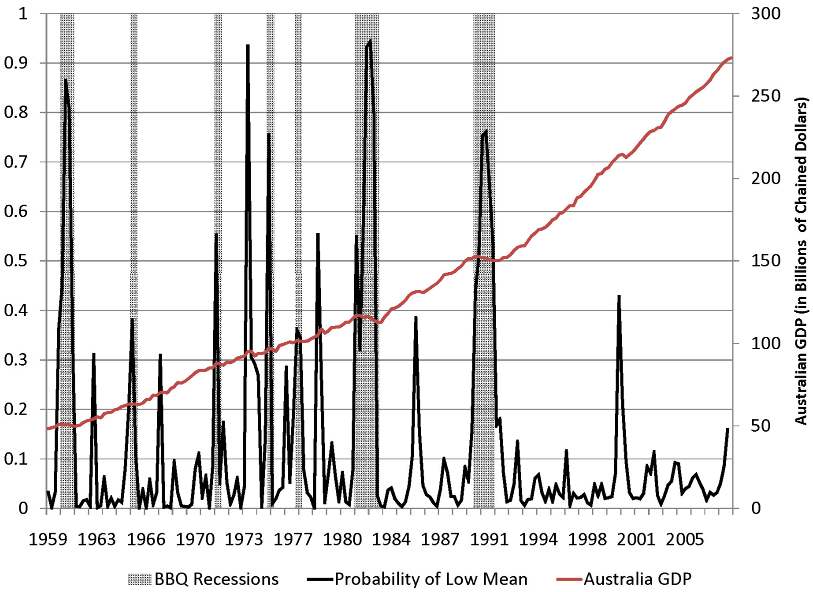


Figure 2.21: Australia Markov-Switching Model with Constant Variance and AR(1) Term: Probabilities of Being in Low-Mean State

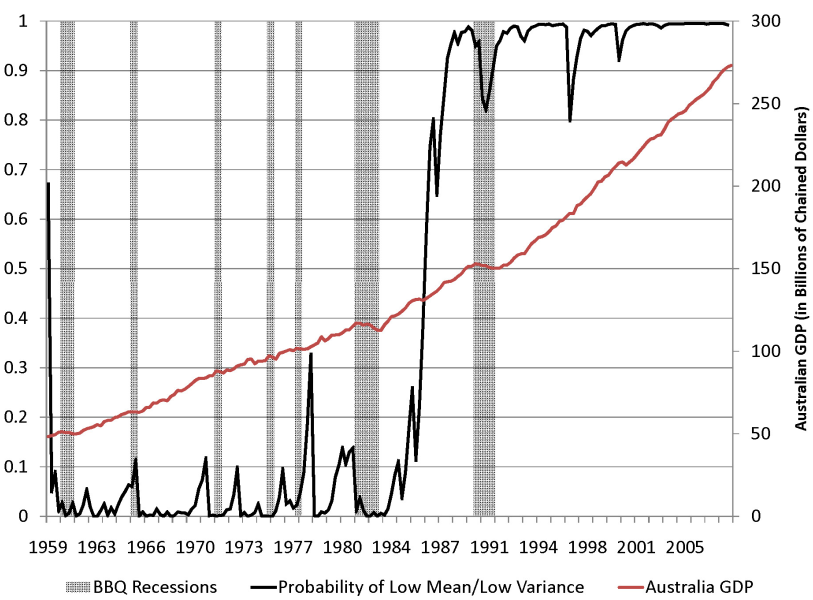


Figure 2.22: Australia Markov-Switching Model with Tied Switching Variance and AR(1) Term: Probabilities of Being in Low-Mean State and Low-Variance State

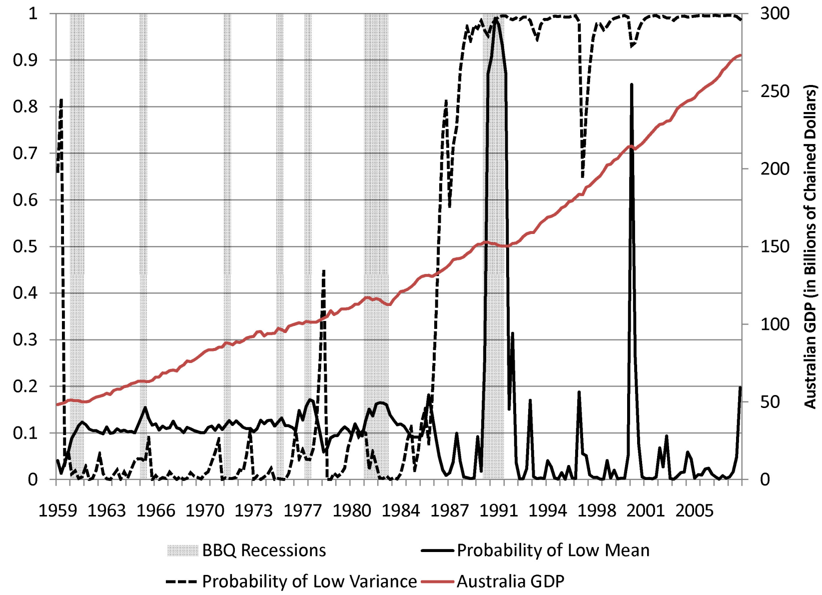


Figure 2.23: Australia Markov-Switching Model with Independent Switching Variance and No AR(1) Term: Probabilities of Being in Low-Mean State and Low-Variance State

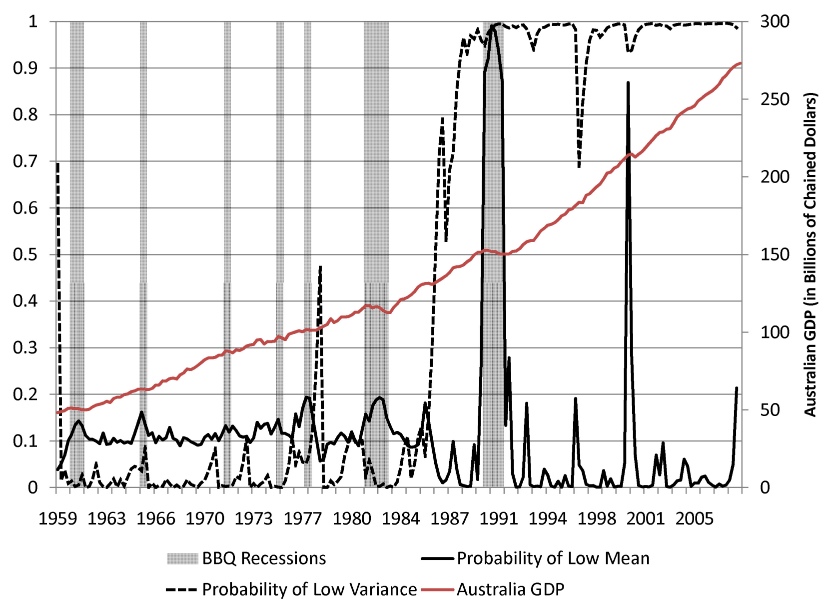


Figure 2.24: Australia Markov-Switching Model with Independent Switching Variance and AR(1) Term: Probabilities of Being in Low-Mean State and Low-Variance State

Chapter 3

Impacts of Changing BBQ Algorithm Parameters on Model Selection

3.1 Introduction

The methods laid out in Chapter Two give us guidance on how to select a model using parametric estimation and non-parametric tests via Monte Carlo simulation. Generally, we found in Chapter Two that the MS-IS-NoAR model is the preferred Markov-switching model for GDP growth for the U.S., U.K., and Australia. Linear models that account for a structural change in volatility also perform well when trying to recreate business cycle features for each of these countries. In this chapter we aim to examine the robustness of non-parametric testing methods. Model selection is determined here in a similar manner as Chapter Two where preference is given to those models that are able to simulate output paths with features statistically close to the various stylized facts described by different algorithms.

3.2 Background on Business Cycle Algorithms

In Harding and Pagan (2003a) the desire to have a transparent business cycle dating method is examined in detail. In referring to the BBQ algorithm that is laid out here, Harding and Pagan state, "...the non-parametric approach seems to be a useful way of constructing business cycle information. It is a very simple algorithm to apply and is very transparent. It is highly robust in that the dates would not change as one changed the sample of observations. It might not of course be robust to major changes in the window width indexed by k." where k is defined as the the number of forward and backward looking quarters in the algorithm. This can be better understood by noting that the algorithm defines a peak as a level which is higher than the previous and subsequent k periods. Additionally, censoring methods are used to ensure that endpoints are not problematic, peaks and troughs alternate, full cycles are at least five quarters, and that more recent peaks are higher than older peaks. The purpose of this chapter is to determine if the models that are preferred are robust to the choice of k as a window width, and the choice of the full cycle length censor m.

Hamilton (2003) is highly critical of the methods of Harding and Pagan (2003a) calling the BBQ algorithm "vague and intuitive" and a "liability rather than an asset." There are a few valid reasons that the BBQ algorithm is open to criticism. Harding and Pagan (2002) note that Burns and Mitchell (1946) interpret a full cycle as having to last between approximately 6 and 15 months, which is agreeable with either a four or five quarter complete cycle. Thus, when Harding and Pagan (2002) note that the BBQ algorithm misses the 1974 recession in the U.K., they use four quarters to define a complete cycle instead of five. This seemingly minor alteration does not impact the cycle statistics for either the U.S. or Australia.[1] Harding and Pagan in fact appear to make a rather arbitrary and vague change so that their model is able to select all pre-determined turning points.

In other work by Morley and Piger (2005) the authors use an optimized algorithm that is designed to match the NBER dating methods more accurately. However, this method does not offer any guidance as to what the algorithm should be calibrated to in any country other than the U.S. or any data series other than GDP. Hess and Iwata (1997) examine the difference between linear and non-linear business cycle models using a much different algorithm. Hess and Iwata use a "running peak" to select the turning points, and do not limit the time that GDP needs to fall below the peak level. Thus, if GDP falls from a peak in time t to a lower level at quarter $t+1$, and then returns to a higher level at $t+2$ relative to t, then a peak is declared at time t and a trough at $t+1$. What is most interesting about the work of Hess and Iwata and Harding and Pagan (2002) is that both articles find that simple linear models and MS models are the only specifications able to reliably recreate business cycle features.

Notably, the 2001 business cycle peak in the U.S. is omitted from the cycle peaks detected in Table 1.2. The standard BBQ algorithm fails to detect a business cycle peak in 2001. A recent article by Leamer (2008) notes the peculiarities of the 2001 business cycle peak in much greater detail. Here we explore the possibility that a different algorithm might be more appropriate. The failure to detect the mild recessions that are experienced under the Great Moderation might be a fatal flaw of using a dating algorithm to determine turning points in the level of GDP. The failure

[1] Harding and Pagan (2002) also study the U.S., U.K., and Australia.

to capture business cycles that may be large declines in economic activity even if only lasting a single quarter becomes a major problem for this algorithm and the models employed here.

A major drawback to using any algorithm to detect turning points is that measurements are only calculated for completed cycles. In the U.S. the standard BBQ algorithm determines that the period between the last detected trough in 1990Q3 until the 2008Q2 peak spanned 70 quarters. Until the most recent peak this very long expansion was not used to calculate the average statistics reported at the bottom of Table 1.2. Information available as of 2009Q1 indicates that the U.S. is in a recessionary state. Until this recession has ended according to the algorithm, the features of the 2008 recession cannot be used to update the stylized facts for recession statistics. For any country and simulation if the final incomplete cycle is appreciably different from the previously estimated cyclical features the estimated facts for that path could change a great deal. The cycle features simulated by a correctly specified model will have increasing difficulty matching an increasingly irrelevant set of statistics since the reference cycle will be calculating averages over two fundamentally different patterns. In general, a historical bias exists in the stylized facts which tends to support more historically aligned models, and not account for the most recent turning points.

3.3 Cyclical Features Under Alternative Algorithms

In order to examine the robustness of the window parameter k for minimum phase length, and the censoring length m for a complete cycle, we merge the techniques of Hess and Iwata (1997) and Harding and Pagan (2002). By examining possible values of $k = 1$ or $k = 2$, and censoring cycles that are either $m = 2, m = 4, m = 5$ quarters in length, we are able to see if the choice of algorithm specification leads to choosing a different model.[2]

The stylized facts will usually change when the parameters of the algorithm change. as is displayed at the bottom of Tables 3.1 - 3.3. In the first two columns of Table 3.1, the stylized facts and peaks of the U.S. business cycle when using the standard BBQ algorithm ($k = 2$ and $m = 5$) are the same as seen throughout Chapter Two. As noted by Harding and Pagan (2002) the U.S. is not impacted by reducing the cycle requirement from $m = 5$ to $m = 4$ in columns 3 and 4. The 2001 recession that is missed by the standard algorithm, is missed due to the fact that the window (k) is too wide and not because the censor (m) is too long. In columns 5 and 6, the window is reduced to $k = 1$ and the cycle length censor is set at $m = 5$. When shortening the window from $k = 2$ to $k = 1$, two additional peaks are detected by the algorithm. Notably, four new peaks are detected, and two that were previously detected are now absent. The peaks of 1956Q2, 1970Q3, 1977Q3, and 2001Q2 are now picked up, while the 1969Q3 and 1980Q1 peaks are missing. The reason that previous peaks are now missing is due to censoring which ensures that peaks and troughs alternate. If we ignored this censoring then it would be unclear how to measure a complete cycle.

In columns 7 and 8 of Table 3.1, the algorithm parameters are set such that $k = 1$ and $m = 4$, which allows the 1980Q3 peak to again be detected. This might seem confusing since the 1980Q1 and 1981Q3 peaks are both detected in the first two algorithms examined here. However, it should be recalled that a peak must be followed by a trough, before a new peak can be detected. The trough following the 1977Q3 peak occurs close enough to the peak to be removed by the censoring

[2]See page 11 for more detail on the definition of a turning point and the construction of the algorithm.

algorithm. Once the entire required cycle length is reduced, the trough isnot removed by the censor and the previously detected peak in 1980Q3 returns. Columns 9 and 10 display an algorithm similar in spirit to the methods of Hess and Iwata who used the running peak method of determining cycles. By setting $k = 1$ and $m = 2$ it is ensured that any time there is even a short downturn, a peak is detected. Likewise, any time declining rates of output reverse and return positive growth, a trough is declared. Using this very sensitive algorithm, 26 peaks are detected in the U.S. post-war period.

While the sensitive algorithm in columns 9 and 10 gives substantially more peaks than the NBER has declared, a representative business cycle model should be able to mimic business cycle characteristics of many different lengths. As the restrictions on the window width and censoring algorithm are reduced, the length and depth of business cycle characteristics reduces dramatically. In comparison to the standard algorithm, recessions using the most sensitive algorithm are estimated as lasting only 1.4 quarters versus 2.9 quarters. Expansions are only estimated to last about 8 quarters (or 2 years) versus the 21.7 (or just under 5.5 years) using the standard algorithm. The implication of using the more sensitive algorithm is that a representative business cycle model should be able to create simulated data that contains short-term drops in level, as well as longer term declines.

For the U.K., moving from the least sensitive algorithms ($k = 2, m = 5$) to one using only a shorter window ($k = 1, m = 5$) leads to an increase in the number of detected and uncensored peaks from 7 to 16. When using $k = 1$ and shortening the complete cycle requirement censor from to $m = 4$, an additional two peaks are detected. Using the least restrictive censor ($k = 1, m = 2$) there are 26 cycle peaks observed in the U.K. data. Downturns using the least restrictive sensor are only 1.6

quarters in length versus 4.2 quarters in the standard algorithm. Expansions are much shorter as well, lasting just 6.9 quarters versus 31 when using the standard algorithm.

In the case of Australia, when using the standard algorithm, the average recession is 3.6 quarters and the typical expansion is 16.5 quarters. By shifting to an algorithm with a shorter phase window ($k = 1, m = 5$) the average recession length declines to 2.1 quarters, and an average expansion length of 9.2 quarters. Using the shortened phase window and censor ($k = 1, m = 4$) yields 16 peaks with average contractions of 2.1 quarters and expansions of 8.5 quarters. Using the least restrictive algorithm ($k = 1, m = 2$) reveals downturns lasting only 1.6 quarters on average and expansions of only 6 quarters. The standard algorithm detects 7 peaks in the Australian GDP series, and 22 peaks using the most liberal algorithm.

Examining the other business cycle measurements for each country shows that moving to a less restrictive algorithm nets shallower contractions (expansions), with less cumulative loss (gain), and more linear fluctuations. These changes are expected. If a single quarter contraction is counted in the statistics, there is no option but for a perfectly linear change. One could also examine the requirement for longer recessions, where $k > 2$ or $m > 5$, but these algorithms tend to pass over entire series in simulation without detecting a complete cycle. In the case of the $k = 2, m = 5$ algorithm, entire simulated series are already missing complete cycles in many cases. Series without cycles typically occur when using a low-volatility measure throughout the entire series. In the case of the MS models, series are sometimes "stuck" by chance in a low-volatility state throughout the simulation, and a complete cycle can not be measured. Nearly all of the conclusions reached using the standard algorithm in Chapter Two hold when including the zeros in the statistics. By examining these less restrictive algorithms, we can test the robustness of our conclusions more stringently.

3.4 Analysis of Algorithm Alternatives

In order to simplify the analysis of the different possible algorithm alternatives, we will examine each country separately in further detail. The simulated data for each of the models that we experiment with are the same as the simulations examined in Chapter Two. Thus, the only difference between any of the results discussed here is the use of a different algorithm to measure cyclical features and censor turning points.

3.4.1 U.S. Modeling Under Alternative Algorithms

Using the standard algorithm ($k = 2, m = 5$) displayed that the best model to replicate the U.S. data was an MS-IS model without an AR(1) coefficient, and that it performed comparably to an ARIMA-SV model where volatility was allowed to shift. The Q-statistics from Table 2.6 show that the features that are created through Monte Carlo simulation of the estimated models are not statistically different than those as measured with the algorithm. Parametric measures of fit do not change for any of these models, and the simulated data are identical to that used in Chapter Two. Therefore discussions about parametric fit are left aside.

A simple change of the full cycle requirement from $m = 5$ to $m = 4$ does not change the measured stylized facts as seen at the bottom of Table 3.1, but a number of simulations very well might be impacted. Examination of Table 3.4 shows that the results differ little from the standard algorithm results in Table 2.6. The MS-IS models still fit the stylized features well, as do the ARIMA-SV models. No other models appear to be any better at fitting the stylized facts. The empirical distributions in

Tables 3.16 and 3.17 also show that the MS-IS models are able to fit all elements of the stylized facts. As noted previously, the MS-IS-AR(1) model is able to simulate all features examined here, but the parametric fit measures show the AR(1) term is insignificant.

When the phase window is shortened from two quarters to one ($k = 2$ to $k = 1$) and the censor is set so a full cycle must be at least five quarters ($m = 5$) the results change dramatically. Examining the Q-statistics in Table 3.7 shows that no model is able to capture all the characteristics of the data. The MS-IS models are able to capture recessionary features with some ability, however, the ability to replicate the updated expansionary characteristics is lost. A closer look at the empirical distributions in Table 3.18 reveals that linear ARIMA-SV models are still able to capture recessionary characteristics. Table 3.19 shows that the MS-IS models are able to capture all but the excess characteristics. The MS-IS models are also somewhat poor at estimating expansionary characteristics with predictions that are too short and shallow.

The third alteration of the standard algorithm that we analyze is where $k = 1$ and $m = 4$, which allows shorter cycles to pass through the original data as well as the simulations. It can be seen in Table 3.20 that no linear model is able to capture more than four characteristics in the simulated data. Table 3.21 shows the MS-IS models perform relatively well from a 90% confidence perspective, both of which are able to capture 7 of 8 features, missing only the excess characteristic of expansions. Table 3.10 presents evidence that the MS-IS-AR(1) model is able to simulate many cyclical features that are observed using this algorithm in the original data. The MS-IS-NoAR model creates cyclical features that are significantly different than those of the underlying data. Considering the very poor fit of the MS-IS-NoAR to the excess feature of expansions and marginal fit of recessionary amplitudes, it is possible that a bounceback MS model might be able to account for these features.

Finally, examining the least sensitive algorithm ($k = 1$ and $m = 2$) in Tables 3.22 and 3.23 shows that the ARIMA-SV and MS-IS models are able to simulate data paths with features that very closely resemble those characteristics that are seen in the original data. Table 3.13 corroborates the findings of the empirical distributions, showing that the Q-tests fail to reject a difference between the features produced in simulated data and actual data for the ARIMA-SV and MS-IS models.

Summarizing the findings by changing each of these algorithms provides significant evidence that the simply specified MS-IS models are capable of replicating nearly all cyclical features observed in the data when using both the standard and very sensitive algorithm parameters. Furthermore, linear models containing a structural change (ARIMA-SV) are also good at simulating features seen in the data using these five different algorithms. These results present even more evidence against the use of MS-CV, or MS-TS models which are able to replicate few of the features observed in the U.S. data for any of these algorithms.

3.4.2 U.K. Modeling Under Alternative Algorithms

The only models able to replicate the cyclical features for the original U.K. data seen in Table 2.7 using the standard algorithm, were an MS-IS-NoAR and a counter-factual structural-change model (SD reversing). The SD reversing specification is designed where the high volatility regime was in place for only 16 years rather than the 37 that the original model estimated. The MS-TS models measured using the standard algorithm are close to being able to capture many features individually as shown in Table 2.10 but are far enough off on most of the features, particularly expansionary amplitudes, to reject the similarities to the actual data. Empirical distributions show in Tables 2.9 and 2.10 that no model is able to simulate paths that match all features seen in the original U.K. data series.

When reducing the cycle censor to $m = 4$ from the standard algorithm, Tables 3.16 and reveals that linear models typically simulate recessions and expansions that are too short and shallow. The MS-TS models for the U.K. are able to fit many features individually as shown in 3.17. Table 3.5 shows that the MS-IS-NoAR model is still able to simulate all features of the underlying data with the new algorithm, albeit every p-value is closer towards rejecting the null of feature similarity at a significant level. As in the case of the standard algorithm, the MS-TS models perform well at projecting recessionary features, but poorly fit expansionary features.

Table 3.8 displays Q-statistics for the algorithm with a phase window $k = 1$ and a cycle censor of $m = 5$. This algorithm gives reasonably good estimates for the U.K. when using the MS-IS models. Examining the individual features when using this algorithm reveals in Table 3.19 that the MS-IS-NoAR model empirical distributions contain all features in a 90% confidence interval. No other linear or MS model is able to replicate all features with this as much regularity as the MS-IS-NoAR model.

The $k = 1, m = 4$ algorithm produces similar results to the $k = 1, m = 5$ algorithm, only emphasizing the improvement of the MS-IS-NoAR estimates in Table 3.21. No linear model captures more than six features. This table also shows that the CV and TS Markov-switching models predict expansions that are far too short. The implication of having expansions that are too short is to estimate too many complete cycles during a single time path. While this statistic is not explicitly calculated here, these numbers imply that too many full cycles are being passed through the algorithm. Q-test results in Table 3.11 reinforce the findings that the MS-IS model is best able to simulate business cycles that are seen in the original data.

Finally, Table 3.23 displays the empirical distribution for the MS models using a $k = 1, m = 2$ algorithm, showing that the IS-NoAR captures 7 of 8 features in a 90% confidence interval. While the expansions predicted by this model are shallower on average than the observed stylized facts, the model performs rather well. Table 3.14 provides further evidence that only the MS-IS models are able to simulate paths that have features that are close to those of the underlying data.

Full examination of the models used to fit U.K. GDP growth reveals that simple linear and structural-change models are not able to recreate business cycle characteristics seen in the original data as the parameters of the algorithm change. These results for the U.K. show that the MS-IS approach is able to replicate both short cycle and standard cyclical features that are observed in the underlying data. The models best at simulating cyclical characteristics of GDP in the U.K. are those that include independent switching volatility. The more convenient constant variance and MS-TS models are unable to replicate business cycle features using any algorithm.

3.4.3 Australia Modeling Under Alternative Algorithms

Recalling the Australian simulation statistics of Table 2.8 shows that several models are able to replicate the observed business cycle features when using the standard algorithm. Including the SD switching model, a random walk, and the MS-IS-NoAR model a high volatility and the counter-factual SD reversing model are also able to replicate all features. Tables 2.9 and 2.10 reinforce the findings from the Q-tests for Australian simulations.

The first algorithm that is examined shortens the cycle censor from $m = 5$ to $m = 4$ while leaving the phase window at $k = 2$. The reduction in the cycle censor helps to rule out a strict linear ARIMA model as seen in Tables 3.16 and 3.17. It can also be seen that the random walk model produces recessions that are too short and shallow. Only the SD switching and SD reversing models are able to match all eight features individually. The MS-IS-NoAR model is able to replicate all features individually, and the Q-tests shown in Table 3.15 fail to reject the similarity between the features of the simulations and original data.

Shortening the phase window to $k = 1$ with a cycle censor of $m = 5$ provides evidence in Table 3.18 that the random walk model is insufficient at simulating data with similar recessionary features to Australian GDP. However, the SD switching and reversing models are still able to create paths with similar features to the original data. Table 3.9 presents Q-tests for this algorithm showing that the strict linear model (GMM No Breaks) is not a good fit. The SD switching, SD reversing, and high volatility models are all able to replicate all features in the aggregate test. It is worth noting that the MS-IS-NoAR poorly fits expansionary features here, predicting recoveries that are far too long according to this algorithm. The actual data shows a downturn approximately every 2.25 years on average, while the MS-IS-NoAR model predicts a downturn once about every 4.75 years on average.

The algorithm using $k = 1, m = 4$ shows that no model is really able to capture all individual features in Tables 3.20 and 3.21. The high volatility, switching volatility, and reversing volatility models are all on the borderline of the 90% confidence interval with some feature. Similarly, marginal success can be seen by the MS-IS-NoAR model. The MS-IS-NoAR model again predicts expansions that are longer than those that are seen when passing the algorithm over the underlying data. However, when examining the Q-tests in Table 3.12 the MS-IS-NoAR model is not rejected at even the 10% level, which is also true for the SD switching model. While there are individual features that are not well explained by these models, they are not rejected due to the structure of the Q-test which weights all features equally in the tests. Individual features that are tested for the MS-IS-NoAR model are likely close enough to the center of the recessionary statistics that they offset the loss in the contractionary states. It is also possible that the spread of the features is not as large for the MS-IS-NoAR model relative to the structural-change SD models, and therefore a smaller loss is calculated into the test statistic.

Finally, the least restrictive $k = 1, m = 2$ algorithm is compared to the underlying data. In Table 3.22 and 3.23 it can be seen that the SD switching and SD reversing models are able to replicate many individual features. These structural-change models perform well at fitting recovery features of Australian data using this algorithm. The IS-NoAR model which performs fairly well when using other algorithms appears to perform rather poorly on expansionary features. However, Table 3.15 provides evidence that the MS-IS-NoAR model is successful at fitting all features of the data, when the SD switching and reversing models are only able to capture features such as amplitude, cumulation, and excess.

3.5 Conclusion

What is clear from testing the algorithm parameters for robustness is that there is not clear evidence supporting any singular model. However, the structural-change model (SD switching) and the MS-IS-NoAR specifications perform very well when changing the parameters of the algorithm. These models are able to create paths that capture both small and large recessions, and short and long expansions. Each of these models has difficulty fitting the asymmetric features of excess and cumulation however which are proxies for the linearity and slope going into and coming out of a trough or peak.

Interesting examples arise from each of the three countries studied here. In the U.S. the MS-IS-NoAR model creates time paths that have features resembling the actual data features when using standard algorithms and also when the algorithm is the least restrictive. Other algorithms are not so forgiving. Across the U.S. model specifications, it is most apparent that the SD switching model is worth considering when trying to determine if a Markov-switching approach is adding anything to the analysis. The most obvious addition to using the MS approach is that we are able to calculate predicted states of the economy from both a volatility and growth perspective. These robustness tests provide further evidence that the MS constant variance and tied switching models are poor at fitting business cycle features.

The results of the U.K. analysis are clearer than for the U.S. or Australia. Here there is only one model, the MS-IS-NoAR, that is able to replicate business cycle features using any algorithm. These results present substantial evidence that the U.K. economy can be modeled very well using an independent switching volatility without a linear structure. Taking the shortcuts of constant or tied volatility will not yield paths with features similar to the underlying data. From a parametric standpoint, the MS-IS-NoAR model has the highest likelihood value in Table 2.4, which may or may not be chosen as the best fit given the non-standard test statistics for an additional nuisance parameter between the TS and IS model.

Paths simulated for Australia show that a linear or structural-change model should be used if approaching the question using a standard algorithm. However, as less restrictive algorithms are applied to the data and simulations, it appears as though the MS-IS-NoAR model is best able to capture the features of the business cycle seen in the underlying data. If additional statistics were included in the test statistics such as the number of complete cycles, our tests might prove otherwise.

Table 3.1: Business Cycle Peaks and Characteristics for U.S. Using Various Phase and Cycle Windows

	Two Q Phase Five Q Cycle		Two Q Phase Four Q Cycle		One Q Phase Five Q Cycle		One Q Phase Four Q Cycle		One Q Phase Two Q Cycle	
	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter
Cycle Peak 1	1948	4	1948	4	1948	4	1948	4	1948	4
Cycle Peak 2	1953	2	1953	2	1953	2	1953	2	1949	3
Cycle Peak 3	1957	3	1957	3	1956	2	1956	2	1953	2
Cycle Peak 4	1960	1	1960	1	1957	3	1957	3	1955	4
Cycle Peak 5	1969	3	1969	3	1960	1	1960	1	1956	2
Cycle Peak 6	1973	4	1973	4	1970	3	1970	3	1957	1
Cycle Peak 7	1980	1	1980	1	1973	4	1973	4	1957	3
Cycle Peak 8	1981	3	1981	3	1977	3	1977	3	1959	2
Cycle Peak 9	1990	3	1990	3	1981	3	1980	1	1960	1
Cycle Peak 10	2008	2	2008	2	1990	3	1981	3	1960	3
Cycle Peak 11					2001	2	1990	3	1969	3
Cycle Peak 12					2008	2	2001	2	1970	3
Cycle Peak 13							2008	2	1973	2
Cycle Peak 14									1973	4
Cycle Peak 15									1974	2
Cycle Peak 16									1977	3
Cycle Peak 17									1980	1
Cycle Peak 18									1981	1
Cycle Peak 19									1981	3
Cycle Peak 20									1982	2
Cycle Peak 21									1990	3
Cycle Peak 22									2000	2
Cycle Peak 23									2000	4
Cycle Peak 24									2001	2
Cycle Peak 25									2007	3
Cycle Peak 26									2008	2
	Cycle Characteristics									
	P→T[a]	T→P[b]	P→T	T→P	P→T	T→P	P→T	T→P	P→T	T→P
Durations	2.8889	21.7000	2.8889	21.7000	2.0909	18.3333	2.0833	16.7692	1.4000	8.0000
Amplitudes(%)	-2.2380	22.1355	-2.2380	22.1355	-1.7135	18.3384	-1.7545	17.0974	-1.0374	8.7365
Cumulative(%)	-2.9379	402.8080	-2.9379	402.8080	-2.0560	259.2070	-2.1465	236.4523	-1.0083	101.1356
Excess(%)	0.0244	1.1160	0.0244	1.1160	0.0597	0.8432	0.0157	0.8263	-0.0139	0.2053

All measures are in quarters
[a] P→T refers to recessionary characteristics.
[b] T→P refers to expansionary characteristics.

Table 3.2: Business Cycle Peaks and Characteristics for U.K. Using Various Phase and Cycle Windows

	Two Q Phase Five Q Cycle		Two Q Phase Four Q Cycle		One Q Phase Five Q Cycle		One Q Phase Four Q Cycle		One Q Phase Two Q Cycle	
	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter
Cycle Peak 1	1955	3	1955	3	1955	3	1955	3	1955	3
Cycle Peak 2	1961	2	1961	2	1958	1	1958	1	1956	1
Cycle Peak 3	1973	2	1973	2	1960	1	1960	1	1957	1
Cycle Peak 4	1974	3	1974	3	1961	2	1961	2	1958	1
Cycle Peak 5	1979	2	1979	2	1964	4	1962	3	1960	1
Cycle Peak 6	1990	2	1990	2	1966	3	1964	4	1961	2
Cycle Peak 7	2008	2	2008	2	1968	1	1966	3	1962	3
Cycle Peak 8					1970	4	1968	1	1964	4
Cycle Peak 9					1973	2	1969	4	1966	3
Cycle Peak 10					1974	3	1970	4	1968	1
Cycle Peak 11					1977	1	1973	2	1969	4
Cycle Peak 12					1979	2	1974	3	1970	4
Cycle Peak 13					1982	2	1977	1	1973	2
Cycle Peak 14					1984	1	1979	2	1974	3
Cycle Peak 15					1990	2	1982	2	1975	1
Cycle Peak 16					2008	2	1984	1	1976	1
Cycle Peak 17							1990	2	1977	1
Cycle Peak 18							2008	2	1978	4
Cycle Peak 19									1979	2
Cycle Peak 20									1979	4
Cycle Peak 21									1981	3
Cycle Peak 22									1982	2
Cycle Peak 23									1984	1
Cycle Peak 24									1990	2
Cycle Peak 25									1992	1
Cycle Peak 26									2008	2
	Cycle Characteristics									
	P→T[a]	T→P[b]	P→T	T→P	P→T	T→P	P→T	T→P	P→T	T→P
Durations	4.1667	31.0000	4.1667	31.0000	2.0667	12.0000	1.9412	10.4706	1.5600	6.8800
Amplitudes(%)	-2.6847	24.2505	-2.6847	24.2505	-1.5835	10.2098	-1.4701	9.0816	-1.1859	6.3617
Cumulative(%)	-7.1992	522.5736	-7.1992	522.5736	-3.0715	142.8478	-2.7466	124.3948	-1.4810	83.3359
Excess(%)	-0.0608	-0.5617	-0.0608	-0.5617	-0.0214	-0.1810	-0.0189	-0.0318	-0.0168	0.0278

All measures are in quarters

[a] P→T refers to recessionary characteristics.

[b] T→P refers to expansionary characteristics.

Table 3.3: Business Cycle Peaks and Characteristics for Australia Using Various Phase and Cycle Windows

	Two Q Phase Five Q Cycle		Two Q Phase Four Q Cycle		One Q Phase Five Q Cycle		One Q Phase Four Q Cycle		One Q Phase Two Q Cycle	
	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter	Year	Quarter
Cycle Peak 1	1960	3	1960	3	1960	3	1960	3	1960	3
Cycle Peak 2	1965	3	1965	3	1963	1	1963	1	1961	1
Cycle Peak 3	1971	3	1971	3	1965	3	1965	3	1963	1
Cycle Peak 4	1975	2	1975	2	1967	4	1967	4	1965	3
Cycle Peak 5	1977	2	1977	2	1970	2	1968	4	1967	1
Cycle Peak 6	1981	3	1981	3	1972	2	1970	2	1967	4
Cycle Peak 7	1990	1	1990	1	1974	1	1972	2	1968	4
Cycle Peak 8					1975	2	1974	1	1970	2
Cycle Peak 9					1977	2	1975	2	1971	3
Cycle Peak 10					1979	1	1977	2	1972	2
Cycle Peak 11					1981	3	1979	1	1974	1
Cycle Peak 12					1986	1	1981	3	1974	3
Cycle Peak 13					1990	1	1986	1	1975	2
Cycle Peak 14					1996	4	1990	1	1976	4
Cycle Peak 15					2000	3	1996	4	1977	2
Cycle Peak 16							2000	3	1979	1
Cycle Peak 17									1981	3
Cycle Peak 18									1982	2
Cycle Peak 19									1986	1
Cycle Peak 20									1990	1
Cycle Peak 21									1996	4
Cycle Peak 22									2000	3
	Cycle Characteristics									
	P→T[a]	T→P[b]	P→T	T→P	P→T	T→P	P→T	T→P	P→T	T→P
Durations	3.5714	16.5000	3.5714	16.5000	2.1333	9.2143	2.0625	8.5333	1.6364	5.9524
Amplitudes(%)	-1.7911	20.0444	-1.7911	20.0444	-1.3768	11.6554	-1.3231	10.9129	-1.1097	7.9494
Cumulative(%)	-3.6672	210.2755	-3.6672	210.2755	-1.9571	66.1284	-1.8510	58.4185	-1.1849	39.1613
Excess(%)	0.0737	0.7977	0.0737	0.7977	0.0231	0.3565	0.0216	0.2856	0.0004	0.1051

All measures are in quarters

[a] P→T refers to recessionary characteristics.

[b] T→P refers to expansionary characteristics.

Table 3.4: Generalized Q-Test Statistics: U.S. Simulation Features vs. Actual Data Features (Two Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.006	0.005	**0.210**	**0.728**	0.006	**0.366**	**0.088**	0.019	**0.108**
GMM No Breaks (2)	0.003	0.002	**0.181**	**0.994**	0.045	**0.905**	**0.066**	0.031	**0.209**
Period 1 Volatility	0.010	**0.632**	0.001	**0.077**	**0.269**	**0.186**	0.004	**0.117**	0.009
Period 2 Volatility	0.000	0.000	**0.396**	**0.350**	0.000	0.000	**0.557**	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.343**	**0.696**	**0.180**	**0.715**	**0.793**	**0.806**	**0.092**	**0.836**	**0.266**
SD Reversing (Vol 2/Vol1)	**0.494**	**0.645**	**0.293**	**0.897**	**0.872**	**0.865**	**0.132**	**0.817**	**0.377**
MS Constant Vol-NoAR	0.000	0.041	0.000	0.010	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.343**	0.000	**0.060**	0.003	0.014	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.099**	0.000	0.011	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	**0.165**	0.000	**0.403**	**0.152**	**0.368**	0.000	0.000	0.002
MS Indep Switching-NoAR	**0.819**	**0.625**	**0.840**	**0.757**	**0.845**	**0.847**	**0.545**	**0.750**	**0.802**
MS Indep Switching-AR(1)	**0.920**	**0.791**	**0.865**	**0.758**	**0.847**	**0.801**	**0.636**	**0.815**	**0.857**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.5: Generalized Q-Test Statistics: U.K. Simulation Features vs. Actual Data Features
(Two Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.010	0.000	0.000	0.000	**0.108**	0.000	0.000
GMM No Breaks (2)	0.000	0.000	**0.152**	0.002	0.000	0.000	**0.176**	0.000	0.000
Period 1 Volatility	0.000	0.000	0.003	0.000	0.000	0.000	**0.121**	0.000	0.000
Period 2 Volatility	0.000	0.000	**0.415**	0.000	0.000	0.000	**0.162**	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.005	0.010	**0.069**	0.007	0.003	0.001	**0.273**	0.006	0.002
SD Reversing (Vol 2/Vol 1)	**0.355**	**0.244**	**0.762**	**0.224**	**0.082**	0.043	**0.708**	**0.235**	**0.150**
MS Constant Vol-NoAR	0.000	**0.471**	0.000	0.000	0.000	0.000	0.002	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.074**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.643**	0.000	**0.550**	0.006	**0.282**	**0.336**	0.000	**0.370**
MS Tied Switching-AR(1)	0.000	**0.719**	0.000	**0.567**	0.017	**0.454**	**0.391**	0.000	**0.464**
MS Indep Switching-NoAR	**0.069**	**0.335**	**0.091**	**0.189**	**0.073**	**0.053**	**0.583**	**0.217**	**0.166**
MS Indep Switching-AR(1)	0.000	**0.684**	0.000	0.026	**0.051**	**0.712**	**0.095**	0.023	**0.242**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.6: Generalized Q-Test Statistics: Australia Simulation Features vs. Actual Data Features (Two Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	**0.380**	**0.277**	**0.505**	**0.112**	**0.770**	**0.372**	**0.303**	**0.221**	**0.259**
GMM No Breaks (2)	**0.059**	0.024	**0.477**	0.038	**0.254**	0.039	**0.411**	**0.065**	0.032
Period 1 Volatility	**0.273**	**0.238**	**0.352**	**0.435**	**0.609**	**0.992**	**0.164**	**0.237**	**0.430**
Period 2 Volatility	0.000	0.000	**0.445**	0.000	0.000	0.000	**0.408**	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.760**	**0.588**	**0.728**	**0.549**	**0.897**	**0.948**	**0.523**	**0.522**	**0.800**
SD Reversing (Vol 2/Vol 1)	**0.774**	**0.641**	**0.726**	**0.538**	**0.929**	**0.930**	**0.397**	**0.555**	**0.752**
MS Constant Vol-NoAR	0.000	**0.060**	0.000	**0.820**	**0.084**	**0.744**	0.039	0.000	**0.104**
MS Constant Vol-AR(1)	0.000	0.000	**0.241**	0.012	0.000	0.000	**0.262**	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.068**	0.000	**0.811**	**0.090**	**0.705**	0.050	0.000	**0.114**
MS Tied Switching-AR(1)	0.000	0.000	**0.354**	0.005	0.000	0.000	**0.055**	0.000	0.000
MS Indep Switching-NoAR	**0.980**	**0.914**	**0.906**	**0.663**	**0.724**	**0.743**	**0.513**	**0.925**	**0.815**
MS Indep Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.574**	0.000	0.000

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.7: Generalized Q-Test Statistics: U.S. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Five Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GMM No Breaks (2)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Period 1 Volatility	0.000	**0.303**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Period 2 Volatility	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.000	**0.101**	0.000	0.000	0.002	0.007	0.001	0.000	0.000
SD Reversing (Vol 2/Vol 1)	0.000	0.014	0.000	**0.176**	0.023	**0.471**	0.000	0.033	0.002
MS Constant Vol-NoAR	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.091**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.064**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Indep Switching-NoAR	0.000	**0.076**	0.011	**0.507**	**0.169**	**0.740**	0.001	0.041	0.007
MS Indep Switching-AR(1)	0.011	**0.180**	**0.137**	**0.873**	**0.557**	**0.999**	0.021	**0.145**	0.101

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.8: Generalized Q-Test Statistics: U.K. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Five Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.169**	0.000	0.000
GMM No Breaks (2)	0.000	0.000	0.000	0.000	0.000	0.000	**0.096**	0.000	0.000
Period 1 Volatility	0.000	0.001	0.000	0.000	0.001	0.000	**0.178**	0.000	0.000
Period 2 Volatility	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.000	0.003	0.000	**0.127**	0.015	0.000	**0.225**	0.036	0.000
SD Reversing (Vol 2/Vol 1)	0.000	0.000	**0.164**	**0.129**	0.000	0.001	**0.435**	0.000	0.005
MS Constant Vol-NoAR	0.000	**0.184**	0.000	0.000	0.000	0.000	**0.128**	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.075**	0.000	0.000	0.000	0.000	**0.090**	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.341**	0.000	0.000	0.000	0.000	**0.064**	0.000	0.000
MS Tied Switching-AR(1)	0.000	**0.382**	0.000	0.000	0.000	0.000	**0.138**	0.000	0.000
MS Indep Switching-NoAR	**0.654**	**0.421**	**0.922**	**0.968**	**0.747**	**0.656**	**0.754**	**0.757**	**0.870**
MS Indep Switching-AR(1)	**0.579**	**0.828**	**0.279**	**0.658**	**0.790**	**0.949**	**0.347**	**0.644**	**0.673**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.9: Generalized Q-Test Statistics: Australia Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Five Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	**0.162**	**0.265**	**0.223**	**0.118**	**0.124**	**0.159**	**0.256**	**0.128**	**0.120**
GMM No Breaks (2)	0.000	0.000	0.035	0.003	0.001	0.000	**0.172**	0.000	0.000
Period 1 Volatility	**0.516**	**0.693**	**0.334**	**0.305**	**0.367**	**0.473**	**0.321**	**0.374**	**0.542**
Period 2 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.164**	**0.458**	**0.157**	**0.238**	**0.721**	**0.687**	**0.384**	**0.110**	**0.524**
SD Reversing (Vol 2/Vol 1)	**0.094**	**0.395**	**0.092**	**0.157**	**0.554**	**0.558**	**0.292**	**0.058**	**0.386**
MS Constant Vol-NoAR	0.000	0.000	0.000	0.163	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.000	0.000	0.008	0.000	0.000	**0.222**	0.000	0.000
MS Tied Switching-NoAR	0.000	0.000	0.000	0.164	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.000	0.001	0.004	0.000	0.000	0.057	0.000	0.000
MS Indep Switching-NoAR	**0.364**	**0.336**	**0.394**	**0.191**	**0.409**	**0.518**	**0.721**	**0.119**	**0.778**
MS Indep Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.780**	0.000	0.000

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.10: Generalized Q-Test Statistics: U.S. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GMM No Breaks (2)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Period 1 Volatility	0.000	**0.161**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Period 2 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.000	0.034	0.000	0.000	0.000	0.002	0.000	0.000	0.000
SD Reversing (Vol 2/Vol 1)	0.000	0.003	0.000	**0.094**	0.004	**0.340**	0.000	0.007	0.000
MS Constant Vol-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-NoAR	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Indep Switching-NoAR	0.000	**0.080**	0.001	**0.316**	**0.060**	**0.561**	0.000	0.012	0.002
MS Indep Switching-AR(1)	0.002	**0.147**	**0.059**	**0.723**	**0.347**	**0.952**	0.014	**0.076**	**0.071**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.11: Generalized Q-Test Statistics: U.K. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.648**	0.000	0.000
GMM No Breaks (2)	0.000	0.000	0.000	0.000	0.000	0.000	**0.502**	0.000	0.000
Period 1 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	**0.654**	0.000	0.000
Period 2 Volatility	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.000	0.000	0.000	0.041	0.008	0.000	**0.678**	0.016	0.000
SD Reversing (Vol 2/Vol 1)	0.000	0.000	**0.211**	**0.112**	0.001	0.000	**0.795**	0.001	0.001
MS Constant Vol-NoAR	0.000	**0.106**	0.000	0.000	0.000	0.000	**0.721**	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.009	0.000	0.000	0.000	0.000	**0.708**	0.000	0.000
MS Tied Switching-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	**0.487**	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.775**	0.000	0.000
MS Indep Switching-NoAR	**0.519**	**0.247**	**0.961**	**0.977**	**0.713**	**0.482**	**0.896**	**0.783**	**0.822**
MS Indep Switching-AR(1)	**0.551**	**0.675**	**0.393**	**0.608**	**0.724**	**0.959**	**0.793**	**0.630**	**0.969**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.12: Generalized Q-Test Statistics: Australia Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Four Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.041	**0.099**	**0.118**	0.024	0.044	0.044	**0.227**	**0.050**	0.033
GMM No Breaks (2)	0.000	0.000	0.021	0.001	0.000	0.000	**0.160**	0.000	0.000
Period 1 Volatility	**0.060**	**0.438**	0.037	0.012	**0.095**	**0.107**	**0.295**	0.019	0.202
Period 2 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.112**	**0.316**	**0.164**	**0.171**	**0.602**	**0.595**	**0.398**	0.076	**0.492**
SD Reversing (Vol 2/Vol 1)	**0.061**	**0.279**	**0.085**	**0.113**	**0.509**	**0.447**	**0.295**	0.037	**0.315**
MS Constant Vol-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.000	0.000	0.008	0.000	0.000	**0.225**	0.000	0.000
MS Tied Switching-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.000	0.001	0.004	0.000	0.000	0.047	0.000	0.000
MS Indep Switching-NoAR	**0.366**	**0.313**	**0.436**	**0.207**	**0.441**	**0.524**	**0.733**	**0.102**	**0.781**
MS Indep Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.694**	0.000	0.000

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.13: Generalized Q-Test Statistics: U.S. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Two Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.001	0.000	0.000	0.000	0.000	0.013	0.000	0.000
GMM No Breaks (2)	0.000	0.001	0.007	**0.469**	0.027	0.030	**0.102**	0.001	0.032
Period 1 Volatility	0.000	**0.527**	0.000	0.000	0.039	0.000	0.063	0.000	0.000
Period 2 Volatility	0.000	0.000	0.003	**0.143**	0.000	0.000	**0.226**	0.000	0.000
SD Switching (Vol 1/Vol 2)	**0.499**	**0.490**	**0.404**	**0.364**	**0.635**	**0.293**	**0.245**	**0.408**	**0.335**
SD Reversing (Vol 2/Vol 1)	**0.152**	**0.264**	**0.126**	**0.788**	**0.710**	**0.987**	**0.174**	**0.120**	**0.468**
MS Constant Vol-NoAR	0.000	**0.342**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Constant Vol-AR(1)	0.000	**0.252**	0.000	0.000	0.000	0.000	0.005	0.000	0.000
MS Tied Switching-NoAR	0.000	**0.788**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Tied Switching-AR(1)	0.000	**0.557**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS Indep Switching-NoAR	**0.632**	**0.565**	**0.624**	**0.931**	**0.774**	**0.982**	**0.248**	**0.589**	**0.593**
MS Indep Switching-AR(1)	**0.710**	**0.594**	**0.734**	**0.670**	**0.746**	**0.898**	**0.499**	**0.506**	**0.803**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.14: Generalized Q-Test Statistics: U.K. Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Two Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.527**	0.000	0.000
GMM No Breaks (2)	0.000	0.000	0.000	0.000	0.000	0.000	**0.341**	0.000	0.000
Period 1 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	**0.538**	0.000	0.000
Period 2 Volatility	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.000	0.001	0.000	0.003	0.001	0.000	**0.647**	0.002	0.000
SD Reversing (Vol 2/Vol 1)	0.001	0.001	**0.089**	0.030	0.001	0.000	**0.620**	0.000	0.002
MS Constant Vol-NoAR	0.000	**0.598**	0.000	0.000	0.000	0.000	**0.463**	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.014	0.000	0.000	0.000	0.000	**0.289**	0.000	0.000
MS Tied Switching-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.026	0.000	0.000
MS Indep Switching-NoAR	**0.831**	**0.698**	**0.960**	**0.795**	**0.457**	**0.327**	**0.831**	**0.789**	**0.680**
MS Indep Switching-AR(1)	**0.722**	**0.887**	**0.525**	**0.690**	**0.659**	**0.972**	**0.726**	**0.827**	**0.945**

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.15: Generalized Q-Test Statistics: Australia Simulation Features vs. Actual Data Features (One Quarter Minimum Phase Length, Two Period Total Cycle Length)

	All Features	Recessions	Expansions	Durations	Amplitudes	Cumulative	Excess	Dur & Amp	Cum & Exc
Random Walk (1)	0.028	0.016	**0.250**	0.001	0.017	0.004	**0.513**	0.004	0.013
GMM No Breaks (2)	0.000	0.000	0.092	0.000	0.000	0.000	**0.518**	0.000	0.000
Period 1 Volatility	0.000	0.023	0.000	0.000	0.007	0.001	**0.470**	0.000	0.003
Period 2 Volatility	0.000	0.000	0.000	0.000	0.000	0.000	**0.791**	0.000	0.000
SD Switching (Vol 1/Vol 2)	0.041	0.040	**0.277**	0.022	**0.461**	**0.419**	**0.676**	0.013	**0.622**
SD Reversing (Vol 2/Vol 1)	0.023	0.038	**0.137**	0.015	**0.439**	**0.274**	**0.530**	0.006	**0.398**
MS Constant Vol-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
MS Constant Vol-AR(1)	0.000	0.000	0.001	0.061	0.000	0.000	**0.887**	0.000	0.000
MS Tied Switching-NoAR	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
MS Tied Switching-AR(1)	0.000	0.000	0.004	0.048	0.000	0.000	**0.490**	0.000	0.000
MS Indep Switching-NoAR	**0.212**	**0.177**	**0.312**	**0.203**	**0.446**	**0.522**	**0.933**	0.036	**0.846**
MS Indep Switching-AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	**0.671**	0.000	0.000

Each cell is a p-value testing the null hypothesis of the actual data features and the simulated features being the same.
A value of zero reflects a firm rejection that the underlying data features and the simulated features are the same.
Failure to reject the null indicates the simulation and underlying data have statistically similar features.
Many simulations produced zero detectable turning points. The simulations where no turning points have been detected are not shown here, and statistics are calculated without those zeros factored in. These results did not affect the general conclusions for any of the models tested here. Results are available upon request.

Table 3.16: Linear Model Simulation Peak to Trough and Trough to Peak Characteristics (Two Quarter Minimum Phase Length, Four Period Total Cycle Length)

		Actual	RW		No Breaks		High Volatility		Low Volatility		Switching Volatility		Reversing Volatility	
U.S.														
P→T	Durations	2.889	2.616	**79%**	2.920	**53%**	3.336	**18%**	2.383	**83%**	3.250	**29%**	3.151	**40%**
	Amplitudes	-2.238	-1.308	0%	-1.399	2%	-2.259	**48%**	-0.579	0%	-2.107	**36%**	-1.913	**26%**
	Cumulation	-2.938	-1.878	**8%**	-2.434	**24%**	-4.729	**85%**	-0.748	1%	-4.399	**72%**	-3.955	**57%**
	Excess	0.024	0.001	**64%**	0.000	**66%**	-0.001	**63%**	-0.001	**70%**	-0.001	**61%**	0.000	**61%**
T→P	Durations	21.700	26.247	**39%**	22.565	**57%**	13.976	96%	64.817	**9%**	17.419	**79%**	24.929	**46%**
	Amplitudes	22.135	24.662	**48%**	21.601	**63%**	15.292	**93%**	50.539	**15%**	17.774	**80%**	23.177	**55%**
	Cumulation	402.808	603.729	**46%**	464.543	**61%**	193.090	**94%**	2709.440	**15%**	319.518	**76%**	675.208	**45%**
	Excess	1.116	0.006	98%	0.004	99%	-0.006	100%	0.003	**87%**	0.090	98%	-0.192	98%
U.K.														
P→T	Durations	4.167	2.959	99%	2.731	99%	2.963	99%	2.085	100%	2.945	97%	2.895	**93%**
	Amplitudes	-2.685	-1.570	0%	-1.182	0%	-1.665	0%	-0.312	0%	-1.644	1%	-1.556	4%
	Cumulation	-7.199	-2.662	0%	-1.790	0%	-2.797	0%	-0.328	0%	-2.783	1%	-2.605	3%
	Excess	-0.061	-0.001	**18%**	0.000	**15%**	0.002	**19%**	0.001	**8%**	-0.002	**25%**	0.004	**28%**
T→P	Durations	31.000	15.747	99%	21.035	**91%**	15.522	99%	72.255	**23%**	16.043	96%	21.443	**80%**
	Amplitudes	24.250	12.821	98%	15.554	**93%**	13.253	98%	45.443	**31%**	13.589	96%	16.575	**81%**
	Cumulation	522.574	179.237	97%	300.489	**89%**	181.884	97%	2440.943	**32%**	197.789	95%	456.248	**79%**
	Excess	-0.562	-0.006	3%	0.003	5%	-0.002	3%	-0.026	**22%**	0.014	**6%**	-0.111	**17%**
Australia														
P→T	Durations	3.571	2.668	96%	2.520	97%	2.949	**89%**	2.176	98%	2.881	**87%**	2.865	**88%**
	Amplitudes	-1.791	-1.565	**25%**	-1.252	**7%**	-2.206	**83%**	-0.575	0%	-2.062	**65%**	-2.011	**62%**
	Cumulation	-3.667	-2.313	**9%**	-1.702	2%	-3.715	**42%**	-0.648	0%	-3.459	**34%**	-3.345	**32%**
	Excess	0.074	-0.001	**81%**	0.000	**83%**	0.000	**78%**	0.000	**86%**	-0.002	**74%**	0.002	**72%**
T→P	Durations	16.500	23.526	**23%**	30.770	**11%**	15.742	**66%**	57.106	**12%**	19.066	**51%**	20.106	**47%**
	Amplitudes	20.044	24.718	**39%**	30.326	**25%**	17.999	**73%**	49.322	**19%**	20.622	**59%**	21.454	**55%**
	Cumulation	210.276	530.714	**22%**	845.180	**13%**	252.138	**56%**	2179.716	**17%**	397.421	**47%**	418.583	**45%**
	Excess	0.798	-0.010	**94%**	0.027	**89%**	0.004	96%	0.020	**81%**	0.071	**89%**	-0.113	**93%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) "RW" is a Random Walk model with constant mean and variance terms.
Column (2) "No Breaks" is an ARIMA(1,1,0) model of output with a constant mean, variance, and AR(1) term.
Columns (3) - (6) are all simulated using estimated parameters from column (3) of Tables 2.1 - 2.3, applying the volatility measures to different parts of the simulations.
Column (3) "High Volatility" is an ARIMA(1,1,0) simulation of output, using the first period variance over the entire time period.
Column (4) "Low Volatility" is an ARIMA(1,1,0) simulation of output, using the second period variance over the entire time period.
Column (5) "Switching Volatility" is an ARIMA(1,1,0) simulation of output, switching the variances (σ_1^2 to σ_2^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.
Column (6) "Reversing Volatility" is an ARIMA(1,1,0) simulation of output, reversing the variances (σ_2^2 to σ_1^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.

Table 3.17: Markov-Switching Simulation Peak to Trough and Trough to Peak Characteristics (Two Quarter Minimum Phase Length, Four Period Total Cycle Length)

		Actual	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
U.S.														
P→T	Durations	2.889	3.755	**5%**	3.220	**25%**	3.809	4%	3.567	**10%**	3.291	**36%**	3.278	**37%**
	Amplitudes	-2.238	-1.920	**16%**	-1.996	**25%**	-2.033	**26%**	-2.096	**33%**	-1.791	**33%**	-1.953	**41%**
	Cumulation	-2.938	-4.476	**83%**	-3.924	**72%**	-4.785	**88%**	-4.690	**89%**	-4.054	**56%**	-4.476	**59%**
	Excess	0.024	0.002	**63%**	-0.008	**69%**	0.003	**62%**	-0.107	**93%**	0.003	**61%**	0.005	**60%**
T→P	Durations	21.700	12.805	98%	13.591	97%	12.924	98%	17.835	**81%**	33.660	**39%**	35.116	**38%**
	Amplitudes	22.135	8.726	100%	11.708	99%	8.950	100%	13.563	95%	31.259	**37%**	36.686	**26%**
	Cumulation	402.808	97.513	100%	141.975	98%	100.665	100%	219.837	**90%**	1104.330	**37%**	1348.307	**32%**
	Excess	1.116	-0.010	100%	-0.005	100%	-0.016	100%	-0.079	100%	0.022	**92%**	-0.025	**91%**
U.K.														
P→T	Durations	4.167	3.927	**69%**	4.015	**63%**	4.371	**43%**	4.360	**43%**	3.166	**93%**	5.048	**69%**
	Amplitudes	-2.685	-2.145	**9%**	-1.965	2%	-2.302	**17%**	-2.337	**19%**	-1.774	3%	-2.242	**16%**
	Cumulation	-7.199	-5.375	**16%**	-4.874	**11%**	-5.734	**20%**	-5.873	**21%**	-3.281	3%	-9.186	**28%**
	Excess	-0.061	0.002	**19%**	0.002	**18%**	-0.022	**36%**	-0.040	**43%**	0.002	**23%**	-0.001	**22%**
T→P	Durations	31.000	9.747	100%	9.573	100%	21.195	**89%**	20.990	**90%**	15.782	**94%**	11.219	99%
	Amplitudes	24.250	6.338	100%	5.254	100%	10.358	98%	10.740	98%	14.174	**93%**	10.043	99%
	Cumulation	522.574	49.821	100%	39.084	100%	197.198	**95%**	207.540	**94%**	300.418	**90%**	132.484	98%
	Excess	-0.562	-0.013	0%	-0.017	0%	-0.060	**5%**	-0.079	**6%**	-0.005	**13%**	-0.007	3%
Australia														
P→T	Durations	3.571	3.590	**55%**	2.367	98%	3.582	**55%**	2.324	98%	4.170	**36%**	6.002	0%
	Amplitudes	-1.791	-1.146	3%	-0.667	1%	-1.143	3%	-0.673	1%	-1.951	**55%**	-7.507	100%
	Cumulation	-3.667	-2.518	**15%**	-0.860	1%	-2.505	**16%**	-0.921	0%	-5.739	**61%**	-34.209	100%
	Excess	0.074	0.001	**90%**	0.000	**92%**	0.002	**90%**	-0.042	97%	0.001	**82%**	0.006	**72%**
T→P	Durations	16.500	25.379	**23%**	45.369	**9%**	25.622	**21%**	48.215	**10%**	28.686	**15%**	8.671	100%
	Amplitudes	20.044	14.173	**86%**	37.048	**23%**	14.259	**86%**	39.318	**23%**	35.269	**16%**	27.140	5%
	Cumulation	210.276	349.542	**48%**	1400.878	**14%**	351.569	**47%**	1557.111	**15%**	925.276	**11%**	169.514	**86%**
	Excess	0.798	-0.012	98%	-0.013	**84%**	-0.009	98%	-0.034	**84%**	-0.280	**89%**	0.222	**84%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) CV-NoAR is a Markov-switching model with a constant variance and no AR(1) term.
Column (2) CV-AR(1) is a Markov-switching model with a constant variance and AR(1) term.
Column (3) TS-NoAR is a Markov-switching model with a tied switching variance and no AR(1) term.
Column (4) TS-AR(1) is a Markov-switching model with a tied switching variance and AR(1) term.
Column (5) IS-NoAR is a Markov-switching model with an independent mean and variance and no AR(1) term.
Column (6) IS-AR(1) is a Markov-switching model with an independent mean and variance with an AR(1) term.

Table 3.18: Linear Model Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Five Period Total Cycle Length)

		Actual	RW		No Breaks		High Volatility		Low Volatility		Switching Volatility		Reversing Volatility	
U.S.														
P→T	Durations	2.091	1.523	98%	1.897	**74%**	2.438	**20%**	1.296	100%	2.157	**47%**	1.955	**67%**
	Amplitudes	-1.713	-0.823	0%	-0.888	0%	-1.636	**39%**	-0.289	0%	-1.309	**12%**	-1.063	4%
	Cumulation	-2.056	-0.827	1%	-1.287	**10%**	-3.061	**79%**	-0.245	0%	-2.386	**53%**	-1.867	**34%**
	Excess	0.060	0.001	99%	0.000	98%	0.000	**89%**	0.000	100%	0.000	**91%**	0.001	**91%**
T→P	Durations	18.333	9.111	100%	10.545	100%	8.577	100%	20.694	**40%**	11.140	100%	13.559	95%
	Amplitudes	18.338	9.439	100%	10.733	100%	9.911	100%	16.896	**69%**	11.358	99%	12.784	96%
	Cumulation	259.207	56.215	100%	82.632	100%	58.597	100%	299.220	**56%**	103.782	99%	151.553	**91%**
	Excess	0.843	0.076	100%	0.077	100%	0.069	100%	0.054	100%	0.103	100%	0.028	100%
U.K.														
P→T	Durations	2.067	1.955	**66%**	1.658	**93%**	1.954	**67%**	1.051	100%	1.828	**78%**	1.527	**93%**
	Amplitudes	-1.584	-1.110	2%	-0.785	0%	-1.189	4%	-0.166	0%	-1.045	2%	-0.713	0%
	Cumulation	-3.071	-1.510	2%	-0.872	0%	-1.591	2%	-0.091	0%	-1.392	2%	-0.879	1%
	Excess	-0.021	0.001	**25%**	0.002	**16%**	0.003	**25%**	0.000	0%	0.001	**27%**	0.002	**21%**
T→P	Durations	12.000	7.985	100%	8.499	100%	7.909	100%	23.793	1%	9.441	96%	14.943	**16%**
	Amplitudes	10.210	7.086	100%	6.928	100%	7.375	99%	15.746	**10%**	8.134	**95%**	11.069	**40%**
	Cumulation	142.848	35.405	100%	37.358	100%	36.226	100%	322.384	**20%**	56.665	99%	152.617	**60%**
	Excess	-0.181	0.055	4%	0.057	2%	0.056	4%	0.024	**12%**	0.067	5%	0.011	**12%**
Australia														
P→T	Durations	2.133	1.594	96%	1.416	100%	1.950	**72%**	1.111	100%	1.656	**92%**	1.617	**93%**
	Amplitudes	-1.377	-1.008	4%	-0.769	0%	-1.574	**73%**	-0.311	0%	-1.143	**19%**	-1.075	**14%**
	Cumulation	-1.957	-1.079	5%	-0.696	0%	-2.114	**50%**	-0.190	0%	-1.453	**20%**	-1.341	**17%**
	Excess	0.023	0.001	**77%**	0.001	**84%**	0.003	**67%**	0.000	99%	0.001	**70%**	0.002	**70%**
T→P	Durations	9.214	8.873	**65%**	9.383	**50%**	7.962	**91%**	15.695	0%	10.350	**25%**	10.779	**18%**
	Amplitudes	11.655	10.257	**84%**	10.223	**83%**	9.932	**90%**	14.530	**21%**	11.290	**63%**	11.597	**57%**
	Cumulation	66.128	58.944	**72%**	63.154	**65%**	49.266	**89%**	178.430	4%	86.478	**33%**	93.350	**28%**
	Excess	0.356	0.085	**93%**	0.081	**95%**	0.076	**92%**	0.070	**93%**	0.087	**90%**	0.047	**93%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) "RW" is a Random Walk model with constant mean and variance terms.
Column (2) "No Breaks" is an ARIMA(1,1,0) model of output with a constant mean, variance, and AR(1) term.
Columns (3) - (6) are all simulated using estimated parameters from column (3) of Tables 2.1 - 2.3, applying the volatility measures to different parts of the simulations.
Column (3) "High Volatility" is an ARIMA(1,1,0) simulation of output, using the first period variance over the entire time period.
Column (4) "Low Volatility" is an ARIMA(1,1,0) simulation of output, using the second period variance over the entire time period.
Column (5) "Switching Volatility" is an ARIMA(1,1,0) simulation of output, switching the variances (σ_1^2 to σ_2^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.
Column (6) "Reversing Volatility" is an ARIMA(1,1,0) simulation of output, reversing the variances (σ_2^2 to σ_1^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.

Table 3.19: Markov-Switching Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Five Period Total Cycle Length)

		Actual	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
		U.S.												
P→T	Durations	2.091	2.562	**15%**	2.370	**25%**	2.585	**15%**	2.283	**35%**	1.953	**62%**	1.971	**60%**
	Amplitudes	-1.713	-1.415	**12%**	-1.441	**18%**	-1.517	**22%**	-1.438	**15%**	-1.026	**12%**	-1.107	**17%**
	Cumulation	-2.056	-2.683	**72%**	-2.576	**66%**	-2.866	**77%**	-2.559	**66%**	-1.891	**39%**	-2.097	**44%**
	Excess	0.060	0.004	**92%**	-0.004	**93%**	0.004	**91%**	-0.047	99%	0.001	**93%**	0.002	**90%**
T→P	Durations	18.333	7.284	100%	8.353	100%	7.312	100%	8.664	100%	14.073	**85%**	16.197	**72%**
	Amplitudes	18.338	5.457	100%	7.619	100%	5.596	100%	7.277	100%	13.937	**90%**	17.603	**66%**
	Cumulation	259.207	24.194	100%	41.959	100%	24.941	100%	39.690	100%	164.883	**87%**	252.655	**70%**
	Excess	0.843	0.027	100%	0.054	100%	0.028	100%	-0.034	100%	0.078	100%	0.084	99%
		U.K.												
P→T	Durations	2.067	3.015	2%	3.003	3%	2.369	**31%**	2.390	**30%**	2.186	**45%**	3.249	**23%**
	Amplitudes	-1.584	-1.743	**68%**	-1.657	**60%**	-1.559	**43%**	-1.576	**46%**	-1.291	**18%**	-1.599	**51%**
	Cumulation	-3.071	-3.757	**66%**	-3.347	**54%**	-2.592	**27%**	-2.700	**32%**	-2.011	**13%**	-3.577	**46%**
	Excess	-0.021	0.003	**30%**	0.007	**27%**	0.004	**28%**	-0.005	**37%**	0.003	**29%**	0.003	**29%**
T→P	Durations	12.000	6.824	100%	6.559	100%	8.272	100%	8.333	100%	10.522	**85%**	7.089	98%
	Amplitudes	10.210	4.763	100%	3.968	100%	4.885	100%	5.100	100%	9.753	**78%**	6.713	96%
	Cumulation	142.848	19.812	100%	15.247	100%	24.743	100%	25.962	100%	165.362	**85%**	54.972	98%
	Excess	-0.181	0.022	2%	0.009	2%	0.032	1%	0.010	3%	0.061	**18%**	0.045	**6%**
		Australia												
P→T	Durations	2.133	1.829	**79%**	1.426	99%	1.817	**79%**	1.373	99%	2.919	**18%**	5.002	0%
	Amplitudes	-1.377	-0.638	0%	-0.356	0%	-0.626	0%	-0.342	0%	-1.354	**43%**	-6.263	100%
	Cumulation	-1.957	-0.934	5%	-0.346	0%	-0.919	4%	-0.339	0%	-3.319	**70%**	-27.511	100%
	Excess	0.023	0.002	**87%**	0.000	**92%**	0.003	**87%**	-0.012	98%	0.000	**68%**	0.005	**58%**
T→P	Durations	9.214	7.905	**94%**	18.971	0%	7.923	**94%**	20.845	0%	18.869	1%	7.886	100%
	Amplitudes	11.655	5.013	100%	16.183	**17%**	5.009	100%	17.795	**13%**	23.488	2%	24.968	0%
	Cumulation	66.128	23.816	100%	255.293	2%	23.811	100%	309.923	2%	391.059	0%	141.910	0%
	Excess	0.356	0.023	100%	0.050	**88%**	0.021	100%	-0.073	**92%**	0.004	**79%**	0.630	**27%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) CV-NoAR is a Markov-switching model with a constant variance and no AR(1) term.
Column (2) CV-AR(1) is a Markov-switching model with a constant variance and AR(1) term.
Column (3) TS-NoAR is a Markov-switching model with a tied switching variance and no AR(1) term.
Column (4) TS-AR(1) is a Markov-switching model with a tied switching variance and AR(1) term.
Column (5) IS-NoAR is a Markov-switching model with an independent mean and variance and no AR(1) term.
Column (6) IS-AR(1) is a Markov-switching model with an independent mean and variance with an AR(1) term.

Table 3.20: Linear Model Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Four Period Total Cycle Length)

		Actual	RW		No Breaks		High Volatility		Low Volatility		Switching Volatility		Reversing Volatility	
U.S.														
P→T	Durations	2.083	1.469	100%	1.794	**85%**	2.215	**37%**	1.285	100%	2.000	**62%**	1.840	**79%**
	Amplitudes	-1.755	-0.798	0%	-0.841	0%	-1.504	**19%**	-0.287	0%	-1.230	**6%**	-1.013	2%
	Cumulation	-2.147	-0.760	0%	-1.125	4%	-2.513	**61%**	-0.239	0%	-2.006	**37%**	-1.598	**21%**
	Excess	0.016	0.000	**78%**	0.000	**74%**	0.000	**66%**	0.000	**93%**	-0.001	**66%**	0.001	**66%**
T→P	Durations	16.769	7.607	100%	9.126	100%	7.172	100%	19.445	**36%**	9.587	100%	11.912	97%
	Amplitudes	17.097	8.151	100%	9.487	100%	8.559	100%	15.943	**68%**	9.992	100%	11.402	97%
	Cumulation	236.452	43.179	100%	68.202	100%	45.586	100%	278.407	**55%**	86.054	99%	129.799	**93%**
	Excess	0.826	0.053	100%	0.045	100%	0.040	100%	0.031	100%	0.071	100%	0.005	100%
U.K.														
P→T	Durations	1.941	1.811	**71%**	1.575	**94%**	1.805	**73%**	1.051	100%	1.710	**80%**	1.471	**95%**
	Amplitudes	-1.470	-1.048	1%	-0.755	0%	-1.124	4%	-0.166	0%	-1.002	2%	-0.707	0%
	Cumulation	-2.747	-1.290	1%	-0.779	0%	-1.359	1%	-0.091	0%	-1.206	1%	-0.798	0%
	Excess	-0.019	0.000	**25%**	0.001	**16%**	0.001	**25%**	0.000	0%	0.000	**27%**	0.001	**23%**
T→P	Durations	10.471	6.455	100%	6.977	100%	6.377	100%	22.673	0%	7.781	98%	12.993	**15%**
	Amplitudes	9.082	6.018	100%	5.926	100%	6.256	100%	15.041	**6%**	6.974	97%	9.802	**41%**
	Cumulation	124.395	25.651	100%	27.756	100%	26.092	100%	305.196	**17%**	43.759	100%	129.558	**62%**
	Excess	-0.032	0.039	**25%**	0.039	**24%**	0.041	**25%**	0.015	**38%**	0.053	**24%**	0.001	**40%**
Australia														
P→T	Durations	2.063	1.527	98%	1.379	100%	1.802	**84%**	1.109	100%	1.578	95%	1.544	96%
	Amplitudes	-1.323	-0.972	4%	-0.752	0%	-1.486	**72%**	-0.311	0%	-1.112	**20%**	-1.047	**14%**
	Cumulation	-1.851	-0.975	3%	-0.651	0%	-1.799	**40%**	-0.189	0%	-1.287	**15%**	-1.194	**12%**
	Excess	0.022	0.000	**79%**	0.000	**86%**	0.001	**70%**	0.000	99%	0.000	**72%**	0.001	**73%**
T→P	Durations	8.533	7.362	**88%**	7.896	**75%**	6.424	99%	14.394	1%	8.718	**49%**	9.136	**37%**
	Amplitudes	10.913	8.827	**93%**	8.869	**92%**	8.427	97%	13.426	**25%**	9.818	**78%**	10.107	**72%**
	Cumulation	58.418	44.790	**84%**	49.155	**76%**	35.571	96%	160.150	4%	68.919	**44%**	75.088	**38%**
	Excess	0.286	0.057	**93%**	0.057	**94%**	0.055	**93%**	0.048	**91%**	0.067	**90%**	0.029	**93%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) "RW" is a Random Walk model with constant mean and variance terms.
Column (2) "No Breaks" is an ARIMA(1,1,0) model of output with a constant mean, variance, and AR(1) term.
Columns (3) - (6) are all simulated using estimated parameters from column (3) of Tables 2.1 - 2.3, applying the volatility measures to different parts of the simulations.
Column (3) "High Volatility" is an ARIMA(1,1,0) simulation of output, using the first period variance over the entire time period.
Column (4) "Low Volatility" is an ARIMA(1,1,0) simulation of output, using the second period variance over the entire time period.
Column (5) "Switching Volatility" is an ARIMA(1,1,0) simulation of output, switching the variances (σ_1^2 to σ_2^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.
Column (6) "Reversing Volatility" is an ARIMA(1,1,0) simulation of output, reversing the variances (σ_2^2 to σ_1^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.

Table 3.21: Markov-Switching Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Four Period Total Cycle Length)

		Actual	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
U.S.														
P→T	Durations	2.083	2.159	**43%**	2.089	**52%**	2.156	**43%**	1.860	**81%**	1.819	**71%**	1.851	**68%**
	Amplitudes	-1.755	-1.261	1%	-1.287	4%	-1.352	3%	-1.249	1%	-0.965	**6%**	-1.043	**10%**
	Cumulation	-2.147	-1.970	**34%**	-1.994	**36%**	-2.083	**40%**	-1.763	**22%**	-1.558	**28%**	-1.758	**34%**
	Excess	0.016	0.001	**69%**	-0.004	**72%**	0.001	**67%**	-0.042	96%	0.001	**72%**	0.002	**69%**
T→P	Durations	16.769	5.365	100%	6.563	100%	5.366	100%	6.167	100%	12.436	**87%**	14.596	**73%**
	Amplitudes	17.097	4.350	100%	6.274	100%	4.452	100%	5.534	100%	12.428	**92%**	15.952	**70%**
	Cumulation	236.452	14.916	100%	29.525	100%	15.309	100%	23.862	100%	142.085	**88%**	224.482	**70%**
	Excess	0.826	0.024	100%	0.032	100%	0.026	100%	-0.067	100%	0.049	100%	0.052	99%
U.K.														
P→T	Durations	1.941	2.494	**5%**	2.465	**7%**	1.673	**86%**	1.692	**84%**	1.979	**48%**	2.806	**25%**
	Amplitudes	-1.470	-1.534	**58%**	-1.482	**50%**	-1.301	**16%**	-1.309	**19%**	-1.201	**17%**	-1.455	**46%**
	Cumulation	-2.747	-2.687	**41%**	-2.401	**27%**	-1.432	1%	-1.490	1%	-1.654	**8%**	-2.801	**37%**
	Excess	-0.019	0.001	**31%**	0.003	**28%**	0.000	**25%**	-0.007	**34%**	0.001	**29%**	0.001	**28%**
T→P	Durations	10.471	5.113	100%	4.822	100%	5.370	100%	5.429	100%	8.570	**87%**	5.752	98%
	Amplitudes	9.082	3.870	100%	3.246	100%	3.504	100%	3.654	100%	8.223	**81%**	5.679	97%
	Cumulation	124.395	12.610	100%	9.423	100%	12.798	100%	13.535	100%	133.929	**85%**	41.036	98%
	Excess	-0.032	0.019	**25%**	0.009	**27%**	0.022	**15%**	0.001	**28%**	0.038	**36%**	0.030	**26%**
Australia														
P→T	Durations	2.063	1.610	**92%**	1.404	99%	1.602	**93%**	1.355	99%	2.744	**17%**	4.997	0%
	Amplitudes	-1.323	-0.585	0%	-0.349	0%	-0.573	0%	-0.337	0%	-1.290	**43%**	-6.258	100%
	Cumulation	-1.851	-0.716	1%	-0.330	0%	-0.702	1%	-0.326	0%	-2.932	**68%**	-27.478	100%
	Excess	0.022	0.001	**93%**	0.000	**92%**	0.001	**93%**	-0.012	98%	0.000	**69%**	0.005	**57%**
T→P	Durations	8.533	5.684	100%	17.439	1%	5.668	100%	19.046	0%	17.469	1%	7.879	100%
	Amplitudes	10.913	3.864	100%	14.970	**19%**	3.843	100%	16.359	**15%**	21.824	2%	24.947	0%
	Cumulation	58.418	14.083	100%	230.766	2%	13.924	100%	279.035	2%	359.206	1%	141.750	0%
	Excess	0.286	0.031	100%	0.019	**87%**	0.029	100%	-0.118	**94%**	-0.010	**78%**	0.629	**21%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) CV-NoAR is a Markov-switching model with a constant variance and no AR(1) term.
Column (2) CV-AR(1) is a Markov-switching model with a constant variance and AR(1) term.
Column (3) TS-NoAR is a Markov-switching model with a tied switching variance and no AR(1) term.
Column (4) TS-AR(1) is a Markov-switching model with a tied switching variance and AR(1) term.
Column (5) IS-NoAR is a Markov-switching model with an independent mean and variance and no AR(1) term.
Column (6) IS-AR(1) is a Markov-switching model with an independent mean and variance with an AR(1) term.

Table 3.22: Linear Model Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Two Period Total Cycle Length)

		Actual	RW		No Breaks		High Volatility		Low Volatility		Switching Volatility		Reversing Volatility	
		U.S.												
P→T	Durations	1.400	1.256	**93%**	1.483	**32%**	1.665	**5%**	1.221	**88%**	1.581	**16%**	1.512	**30%**
	Amplitudes	-1.037	-0.690	0%	-0.705	2%	-1.167	**73%**	-0.271	0%	-0.999	**39%**	-0.855	**19%**
	Cumulation	-1.008	-0.522	0%	-0.752	**16%**	-1.483	**83%**	-0.207	0%	-1.246	**62%**	-1.041	**42%**
	Excess	-0.014	0.000	**13%**	0.000	**20%**	0.000	**30%**	0.000	**7%**	0.000	**30%**	0.000	**29%**
T→P	Durations	8.000	4.890	100%	6.624	**87%**	4.775	100%	16.884	0%	6.740	**85%**	8.793	**36%**
	Amplitudes	8.736	5.671	100%	7.128	**86%**	6.013	98%	13.951	**7%**	7.300	**84%**	8.642	**58%**
	Cumulation	101.136	24.877	100%	47.203	98%	27.850	100%	240.977	**12%**	58.028	**93%**	93.124	**68%**
	Excess	0.205	0.001	100%	0.001	98%	0.000	99%	-0.002	**88%**	0.022	**95%**	-0.027	97%
		U.K.												
P→T	Durations	1.560	1.373	**95%**	1.287	99%	1.360	96%	1.045	100%	1.332	96%	1.251	97%
	Amplitudes	-1.186	-0.828	0%	-0.631	0%	-0.886	1%	-0.165	0%	-0.823	1%	-0.638	0%
	Cumulation	-1.481	-0.724	0%	-0.494	0%	-0.757	0%	-0.090	0%	-0.699	0%	-0.527	1%
	Excess	-0.017	0.000	**15%**	0.000	**8%**	0.000	**15%**	0.000	0%	0.000	**17%**	0.000	**17%**
T→P	Durations	6.880	3.682	100%	4.169	100%	3.568	100%	20.488	0%	4.637	99%	8.905	**12%**
	Amplitudes	6.362	3.869	100%	3.933	100%	3.980	100%	13.666	1%	4.582	98%	7.026	**36%**
	Cumulation	83.336	12.324	100%	14.336	100%	12.139	100%	274.480	**7%**	23.909	100%	86.256	**61%**
	Excess	0.028	0.001	**68%**	0.001	**70%**	0.000	**68%**	-0.005	**59%**	0.010	**62%**	-0.022	**70%**
		Australia												
P→T	Durations	1.636	1.277	100%	1.217	100%	1.362	99%	1.086	100%	1.296	99%	1.280	99%
	Amplitudes	-1.110	-0.828	2%	-0.665	0%	-1.174	**63%**	-0.302	0%	-0.956	**19%**	-0.910	**13%**
	Cumulation	-1.185	-0.644	1%	-0.474	0%	-1.007	**22%**	-0.177	0%	-0.804	**10%**	-0.754	**7%**
	Excess	0.000	0.000	**52%**	0.000	**52%**	0.000	**51%**	0.000	**67%**	0.000	**51%**	0.000	**50%**
T→P	Durations	5.952	4.617	**95%**	5.199	**81%**	3.629	100%	11.987	0%	5.575	**69%**	5.938	**55%**
	Amplitudes	7.949	6.038	96%	6.282	**92%**	5.387	99%	11.361	**10%**	6.768	**84%**	7.031	**79%**
	Cumulation	39.161	24.899	**92%**	29.314	**83%**	16.812	99%	130.558	2%	40.982	**57%**	45.715	**48%**
	Excess	0.105	0.000	**88%**	0.001	**88%**	0.002	**89%**	0.005	**77%**	0.011	**82%**	-0.015	**88%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) "RW" is a Random Walk model with constant mean and variance terms.
Column (2) "No Breaks" is an ARIMA(1,1,0) model of output with a constant mean, variance, and AR(1) term.
Columns (3) - (6) are all simulated using estimated parameters from column (3) of Tables 2.1 - 2.3, applying the volatility measures to different parts of the simulations.
Column (3) "High Volatility" is an ARIMA(1,1,0) simulation of output, using the first period variance over the entire time period.
Column (4) "Low Volatility" is an ARIMA(1,1,0) simulation of output, using the second period variance over the entire time period.
Column (5) "Switching Volatility" is an ARIMA(1,1,0) simulation of output, switching the variances (σ_1^2 to σ_2^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.
Column (6) "Reversing Volatility" is an ARIMA(1,1,0) simulation of output, reversing the variances (σ_2^2 to σ_1^2) at the pre-determined break point from the last row in column (3) of Tables 2.1 - 2.3.

Table 3.23: Markov-Switching Simulation Peak to Trough and Trough to Peak Characteristics (One Quarter Minimum Phase Length, Two Period Total Cycle Length)

		Actual	CV-NoAR		CV-AR(1)		TS-NoAR		TS-AR(1)		IS-NoAR		IS-AR(1)	
U.S.														
P→T	Durations	1.400	1.398	**53%**	1.595	**9%**	1.385	**59%**	1.288	**85%**	1.437	**47%**	1.505	**38%**
	Amplitudes	-1.037	-0.898	**10%**	-0.999	**39%**	-0.960	**24%**	-0.931	**19%**	-0.783	**27%**	-0.859	**37%**
	Cumulation	-1.008	-0.863	**23%**	-1.198	**65%**	-0.898	**27%**	-0.856	**24%**	-0.898	**38%**	-1.094	**49%**
	Excess	-0.014	0.000	**15%**	-0.003	**30%**	0.000	**16%**	-0.029	**77%**	0.000	**22%**	0.001	**26%**
T→P	Durations	8.000	2.600	100%	4.376	100%	2.615	100%	3.899	100%	8.897	**51%**	11.174	**35%**
	Amplitudes	8.736	2.542	100%	4.446	100%	2.612	100%	3.703	100%	9.055	**54%**	12.308	**19%**
	Cumulation	101.136	5.574	100%	17.918	100%	5.795	100%	13.716	100%	98.963	**68%**	169.867	**43%**
	Excess	0.205	-0.001	100%	0.001	100%	-0.001	100%	-0.067	100%	0.001	**95%**	0.000	**89%**
U.K.														
P→T	Durations	1.560	1.509	**66%**	1.392	**93%**	1.023	100%	1.028	100%	1.429	**80%**	1.886	**51%**
	Amplitudes	-1.186	-1.043	**14%**	-1.003	4%	-0.916	1%	-0.919	1%	-0.919	**8%**	-1.076	**26%**
	Cumulation	-1.481	-1.104	**11%**	-0.877	1%	-0.475	0%	-0.489	0%	-0.869	5%	-1.481	**32%**
	Excess	-0.017	0.000	**17%**	0.000	**12%**	0.000	1%	-0.002	3%	0.000	**21%**	0.000	**20%**
T→P	Durations	6.880	2.496	100%	2.029	100%	2.969	100%	3.075	100%	4.979	**89%**	3.417	98%
	Amplitudes	6.362	2.275	100%	1.831	100%	2.167	100%	2.270	100%	5.196	**87%**	3.683	98%
	Cumulation	83.336	4.649	100%	2.683	100%	6.204	100%	6.765	100%	80.740	**87%**	24.422	98%
	Excess	0.028	0.000	**79%**	0.000	**85%**	0.000	**88%**	-0.013	**94%**	-0.003	**60%**	-0.001	**71%**
Australia														
P→T	Durations	1.636	1.226	100%	1.338	**92%**	1.224	100%	1.291	95%	2.237	**11%**	4.523	0%
	Amplitudes	-1.110	-0.468	0%	-0.331	0%	-0.460	0%	-0.318	0%	-1.091	**43%**	-5.674	100%
	Cumulation	-1.185	-0.379	0%	-0.292	0%	-0.373	0%	-0.290	0%	-2.117	**75%**	-24.338	100%
	Excess	0.000	0.000	**53%**	0.000	**54%**	0.000	**53%**	-0.011	**83%**	0.000	**51%**	0.002	**49%**
T→P	Durations	5.952	2.776	100%	15.720	0%	2.777	100%	17.292	0%	13.993	0%	6.794	0%
	Amplitudes	7.949	2.269	100%	13.571	**5%**	2.263	100%	14.915	4%	17.557	2%	22.019	0%
	Cumulation	39.161	5.456	100%	206.537	1%	5.420	100%	252.163	1%	284.805	0%	119.781	0%
	Excess	0.105	0.000	100%	-0.003	**71%**	0.000	100%	-0.124	**85%**	-0.027	**66%**	0.439	**18%**

Bold percentages indicate that the simulation value is within a 90% confidence interval.
Column (1) CV-NoAR is a Markov-switching model with a constant variance and no AR(1) term.
Column (2) CV-AR(1) is a Markov-switching model with a constant variance and AR(1) term.
Column (3) TS-NoAR is a Markov-switching model with a tied switching variance and no AR(1) term.
Column (4) TS-AR(1) is a Markov-switching model with a tied switching variance and AR(1) term.
Column (5) IS-NoAR is a Markov-switching model with an independent mean and variance and no AR(1) term.
Column (6) IS-AR(1) is a Markov-switching model with an independent mean and variance with an AR(1) term.

Chapter 4

Summary of Findings

It is apparent throughout the simulation experiments performed in Chapter Two, and the tests performed in Chapter Three, that one should not be using a simple linear or random walk structure to justify using a Markov-switching model. Many authors have modeled the business cycle using constant variance or tied variance approaches to the MS model, but almost universally these authors justify their use by showing how much better their estimations and simulations are relative to a simple linear model. The Great Moderation has changed this. Now to replicate the business cycle features of an economy, you must properly account for the shift in volatility that occurred in the mid-1980's and early-1990's.

The robustness checks of Chapter Three show that using these algorithms has substantial merit, as many authors have questioned their use. These results lay the groundwork for future examination of the different types of cycles that are created by different models. It should be clear from these robustness checks that the choice of model is not necessarily dependent upon the parameters used for the algorithm checking the data. In most cases the different algorithm parameters select the same model over and over again. In a few cases, algorithms do not find any preferred model. This provides hope that some other model might be able to estimate and simulate data that replicates all features of a data series regardless of the parameters of the algorithm used to measure its success. Specifically it would be interesting to evaluate a number of adapted MS models against one another, such as the bounceback model proposed by Kim, Morley, and Piger (2005). It is also worth noting that researchers should examine all features of the data that is being simulated, since many models are unable to capture both recessionary and expansionary features.

In nearly all of the experiments performed here, the Markov-switching model with independently switching mean and volatility is the preferred specification. The most attractive feature of the MS-IS model is that we are able to account for the Great Moderation, as well as the possibility that it is ending. The evidence presented in Chapter Two shows substantial new evidence that the Great Moderation has ended in the U.K., and possibly is coming to an end in the U.S. as well. If the Great Moderation has come to a conclusion, many of the leading theories for its causes should be reexamined. For example, if the primary cause of the Great Moderation was better monetary policy, then what has caused this to suddenly change so dramatically. If the cause has been the lack of large shocks over the past 25 years, why have larger shocks all of a sudden returned. Each of these findings is important to understanding the cause and impact of business cycles which have again taken a prominent role in macroeconomic research.

Appendix A

Testing and Estimation Procedures

It should be pointed out that statistical tests moving from a linear, to a linear model with volatility break are tested here. However, tests moving from the linear model with break to the MS model are not performed here. Hansen (1992) and Garcia (1998) both cite valid inference of MS models versus linear alternatives using Likelihood Ratio testing procedures. However, the test statistics and p-values must be calculated for each model that is tested here. Creation of the LR test statistic faces a number of difficulties cited by Hansen (1992) and Garcia (1998). The clearest hurdle to overcome with regard to this study is noted by Carrasco (2002) who shows that tests for a simple structural-change model are not valid if the correct model is that of a threshold autoregression (TAR) or MS type. Carrasco (2002) notes that her results do not necessarily extend easily to a model where the volatility is tested to switch in the AR specification, or where the MS approach has a similar nuisance parameter. Previous work by Kim, Morley, and Piger (2005) shows the computational barriers to testing for the presence of Markov-switching over a null linear model with no switching. Kim, Morley, and Piger (2005) are able to reject the null of a linear model in favor of the Markov-switching approach. However, they do not show any tests for validity when moving from a MS model without switching variance, to a model with an additional nuisance parameter. The creation of a separate test for each country moving from the linear to MS models is not performed here, as it is beyond the scope of this paper. This paper aims to display the statistical difference between simulated data across models at fitting features of the business cycle. The statistical validity of moving from a linear to a non-linear model is important, and is definitely in need of further study.

Constant volatility and tied switching Markov-switching models are estimated using James Hamilton's publicly provided numerical estimation methods at
http://weber.ucsd.edu/ jhamilto/software.html using OxGauss. Independent switching models are estimated using Maximo Camacho and Gabriel Perez-Quiros publicly provided numerical estimation methods at
http://www.bepress.com/snde/vol11/iss4/art3/ using OxGauss.

Appendix B

Appendix Figures and Tables

The following tables are tests of the linear model specification chosen to be used for the linear and structural-change models.

Table A-1: Specification Tests for the U.S.

	Test All Parameters for Breaks					Drop C		Drop C&SD
	C	AR	C & AR	AR & SD	SD	SD	AR	AR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant (Pre-Break)	0.6973	0.3737	0.5694	0.4193	0.5595	0.5407	0.5247	0.5493
(P-values in Parentheses)	(0.000)	(0.002)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Constant (Post-Break)	0.4638	0.5520	0.4438	0.5686	0.5089			
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)			
AR(1) (Pre-Break)	0.2699	0.5308	0.3255	0.3900	0.3197	0.3353	0.4983	0.3438
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AR(1) (Post-Break)	0.3270	0.2937	0.1951	0.2972	0.2783	0.2555	0.3057	-0.0106
	(0.000)	(0.000)	(0.079)	(0.000)	(0.005)	(0.000)	(0.000)	(0.906)
Std. Deviation (Pre-Break)	1.0088	1.0018	0.8746	1.2279	1.0823	1.0822	1.0161	0.8276
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)	0.6819	0.7903	0.5155	0.7301	0.4877	0.4816	0.7943	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Wald Test Values								
Sup	2.6027	10.4553	7.3335	20.1076	45.7916	54.3560	11.1808	18.8582
	(0.947)	(0.082)	(0.701)	(0.012)	(0.000)	(0.000)	(0.061)	(0.002)
Ave	0.3654	1.4379	1.1501	7.1011	19.0459	23.3745	2.0203	4.8637
	(0.963)	(0.333)	(0.926)	(0.008)	(0.000)	(0.000)	(0.611)	(0.065)
Exp	0.6286	0.9201	1.7080	9.1076	14.5084	17.0112	2.3864	5.0764
	(0.953)	(0.825)	(0.963)	(0.017)	(0.000)	(0.000)	(0.822)	(0.230)
J-Statistic	2.4889	1.1246	1.9367	1.2680	1.9355	1.9800	1.7003	3.9285
Break Dates								
Wald Statistic	1973Q4	1956Q3	1998Q4	1960Q4	1984Q1	1984Q1	1956Q2	1998Q4

Values are calculated for these parameters when allowing for breaks in all parameters. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the constant represents the constant before the break. The second reported value for each parameter estimates the value after the break.

*Wald tests here signify a test of the restriction rather than a test for the break date.

The parameter values are those from the estimated break date supported by the model.

Column (1) uses a Wald test of the restriction that the constants do not change. The Wald test value is from the most likely period in which this would occur

Column (2) uses a Wald test of the restriction that the AR coefficients do not change.

Column (3) uses a Wald test of the restriction that the AR coefficients AND constants do not change.

Column (4) uses a Wald test of the restriction that the AR coefficients AND standard deviations do not change.

Column (5) uses a Wald test of the restriction that the standard deviations do not change. This is similar to the tests performed in Tables 1-3.

Column (6) uses a Wald test of the restriction that the standard deviations do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.

Column (7) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.

Column (8) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in either the constant or the standard deviations.

Table A-2: Specification Tests for the U.K.

	Test All Parameters for Breaks					Drop C		Drop C&SD
	C	AR	C & AR	AR & SD	SD	SD	AR	AR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant (Pre-Break)	0.6581	0.6580	0.6571	0.6581	0.6150	0.5763	0.5807	0.5810
(P-values in Parentheses)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant (Post-Break)	-0.3439	-0.3535	-0.3332	-0.3439	0.4241			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
AR(1) (Pre-Break)	-0.1043	-0.1040	-0.1042	-0.1043	-0.1057	-0.0969	-0.0851	-0.0859
	(0.181)	(0.182)	(0.181)	(0.181)	(0.181)	(0.210)	(0.284)	(0.262)
AR(1) (Post-Break)	1.4080	1.4169	1.4078	1.4080	0.4063	0.2078	0.2267	0.6442
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.146)	(0.008)	(0.000)
Std. Deviation (Pre-Break)	0.9752	0.9778	0.9772	0.9752	1.0815	1.1131	1.1858	0.9449
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)	0.3752	0.3646	0.3820	0.3752	0.3083	0.2850	0.4504	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Wald Test Values								
Sup	59.3718	203.8519	28.7147	38.1886	64.6291	76.7353	15.4635	35.7081
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)	(0.000)
Ave	25.9435	96.9596	10.8953	15.7754	28.1135	34.3588	4.8038	13.2736
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.068)	(0.000)
Exp	8.6196	24.8658	5.5663	3.9968	24.2492	30.4935	5.1251	10.2930
	(0.003)	(0.000)	(0.174)	(0.415)	(0.000)	(0.000)	(0.225)	(0.008)
J-Statistic	2.3777	2.5016	2.3881	2.3777	0.8568	3.3433	3.8317	7.0781
Break Dates								
Wald Statistic	1999Q2	1999Q3	1999Q1	1999Q2	1992Q2	1991Q3	1979Q3	1990Q3

Values are calculated for these parameters when allowing for breaks in all parameters. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the constant represents the constant before the break. The second reported value for each parameter estimates the value after the break.

*Wald tests here signify a test of the restriction rather than a test for the break date.

The parameter values are those from the estimated break date supported by the model.

Column (1) uses a Wald test of the restriction that the constants do not change. The Wald test value is from the most likely period in which this would occur

Column (2) uses a Wald test of the restriction that the AR coefficients do not change.

Column (3) uses a Wald test of the restriction that the AR coefficients AND constants do not change.

Column (4) uses a Wald test of the restriction that the AR coefficients AND standard deviations do not change.

Column (5) uses a Wald test of the restriction that the standard deviations do not change. This is similar to the tests performed in Tables 1-3.

Column (6) uses a Wald test of the restriction that the standard deviations do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.

Column (7) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.

Column (8) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in either the constant or the standard deviations.

Table A-3: Specification Tests for Australia

	Test All Parameters for Breaks					Drop C		Drop C&SD
	C	AR	C & AR	AR & SD	SD	SD	AR	AR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant (Pre-Break)	1.1358	1.1105	1.3884	0.8341	1.0026	0.8927	0.8649	0.7012
(P-values in Parentheses)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant (Post-Break)	0.4133	0.4277	0.7853	0.8831	0.4821			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)			
AR(1) (Pre-Break)	-0.1964	-0.1889	-0.1024	0.1809	-0.1214	-0.0545	0.1591	-0.0364
	(0.006)	(0.016)	(0.528)	(0.030)	(0.239)	(0.549)	(0.027)	(0.675)
AR(1) (Post-Break)	0.4471	0.4382	-0.0068	-0.0390	0.3903	0.0196	-0.0715	0.4136
	(0.000)	(0.000)	(0.960)	(0.728)	(0.004)	(0.868)	(0.551)	(0.000)
Std. Deviation (Pre-Break)	1.3961	1.4002	1.4392	1.5444	1.4141	1.4134	1.5414	1.1321
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. Deviation (Post-Break)	0.6211	0.6219	0.8754	0.9469	0.5756	0.5513	0.8841	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Wald Test Values								
Sup	14.9749	38.0462	6.4072	21.4919	100.2612	116.9350	5.6023	24.0986
	(0.011)	(0.000)	(0.791)	(0.006)	(0.000)	(0.000)	(0.460)	(0.000)
Ave	3.4837	14.1166	0.9819	6.0024	45.4600	53.5927	0.8988	8.8084
	(0.037)	(0.000)	(0.961)	(0.023)	(0.000)	(0.000)	(0.978)	(0.002)
Exp	3.0956	4.9743	1.0663	4.8861	39.4814	44.3717	1.1641	6.5347
	(0.165)	(0.039)	(1.000)	(0.258)	(0.000)	(0.000)	(1.000)	(0.097)
J-Statistic	4.5126	4.4640	3.3601	1.9130	5.1866	4.1007	2.8749	11.7638
Break Dates								
Wald Statistic	1982Q3	1982Q4	1971Q3	1967Q1	1985Q2	1985Q2	1970Q1	1983Q4

Values are calculated for these parameters when allowing for breaks in all parameters. Parameter values are those from the most likely occurring break date. The first reported value for each parameter when allowing for breaks, such as the constant represents the constant before the break. The second reported value for each parameter estimates the value after the break.
*Wald tests here signify a test of the restriction rather than a test for the break date.
The parameter values are those from the estimated break date supported by the model.
Column (1) uses a Wald test of the restriction that the constants do not change. The Wald test value is from the most likely period in which this would occur
Column (2) uses a Wald test of the restriction that the AR coefficients do not change.
Column (3) uses a Wald test of the restriction that the AR coefficients AND constants do not change.
Column (4) uses a Wald test of the restriction that the AR coefficients AND standard deviations do not change.
Column (5) uses a Wald test of the restriction that the standard deviations do not change. This is similar to the tests performed in Tables 1-3.
Column (6) uses a Wald test of the restriction that the standard deviations do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.
Column (7) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in the constant and is is similar to the tests performed in Tables 1-3.
Column (8) uses a Wald test of the restriction that the AR coefficients do not change. This model does not include a break in either the constant or the standard deviations.

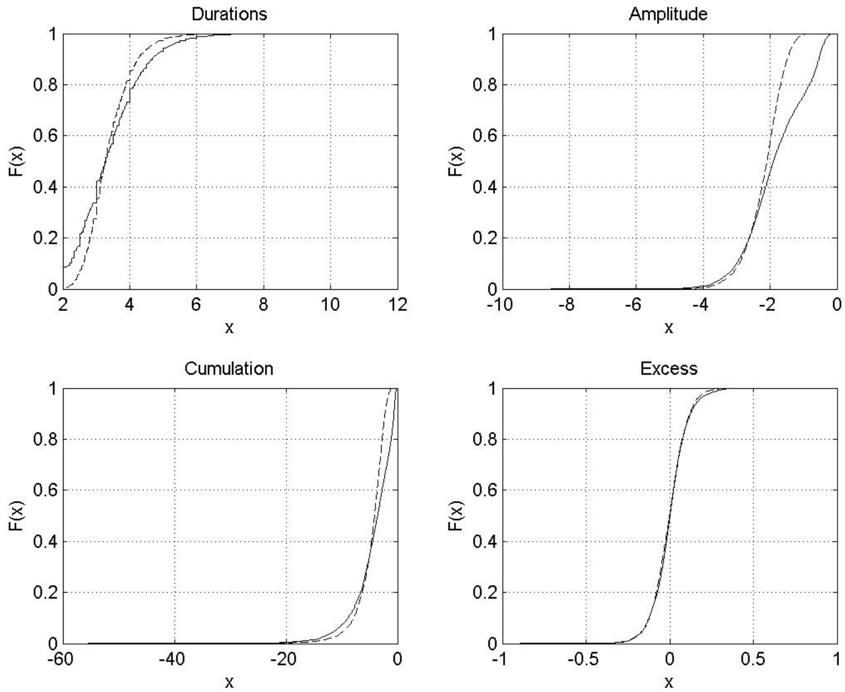


Figure A-1: U.S. Empirical Distributions of Average Recession Characteristics
ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

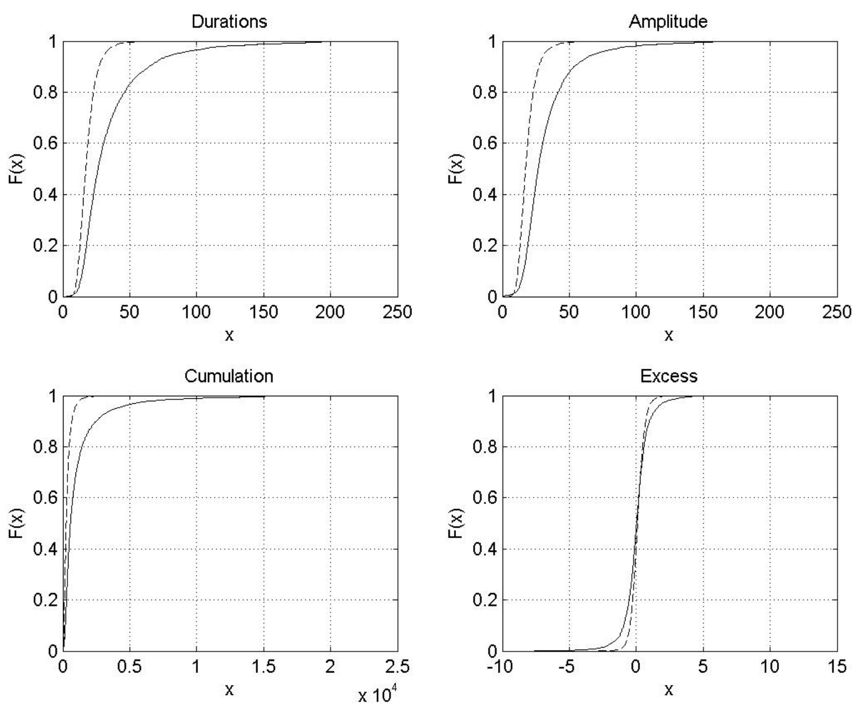


Figure A-2: U.S. Empirical Distributions of Average Expansion Characteristics
ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

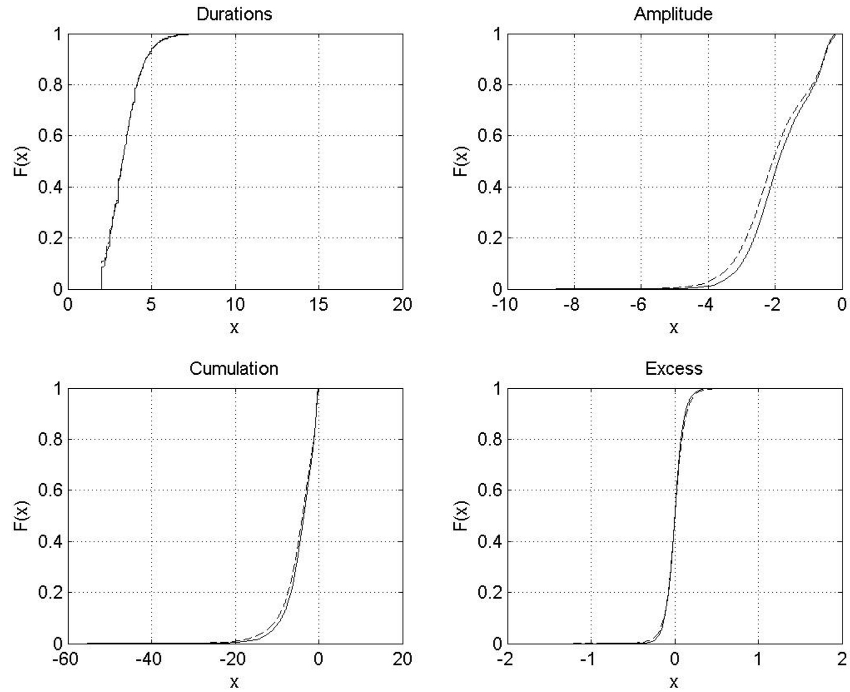


Figure A-3: U.S. Empirical Distributions of Average Recession Characteristics MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

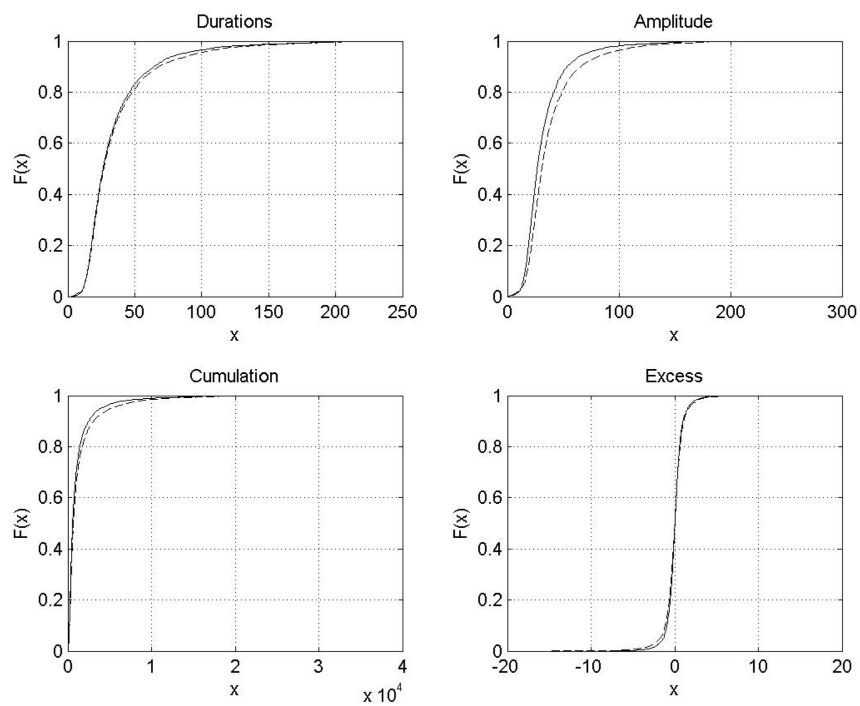


Figure A-4: U.S. Empirical Distributions of Average Expansion Characteristics MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

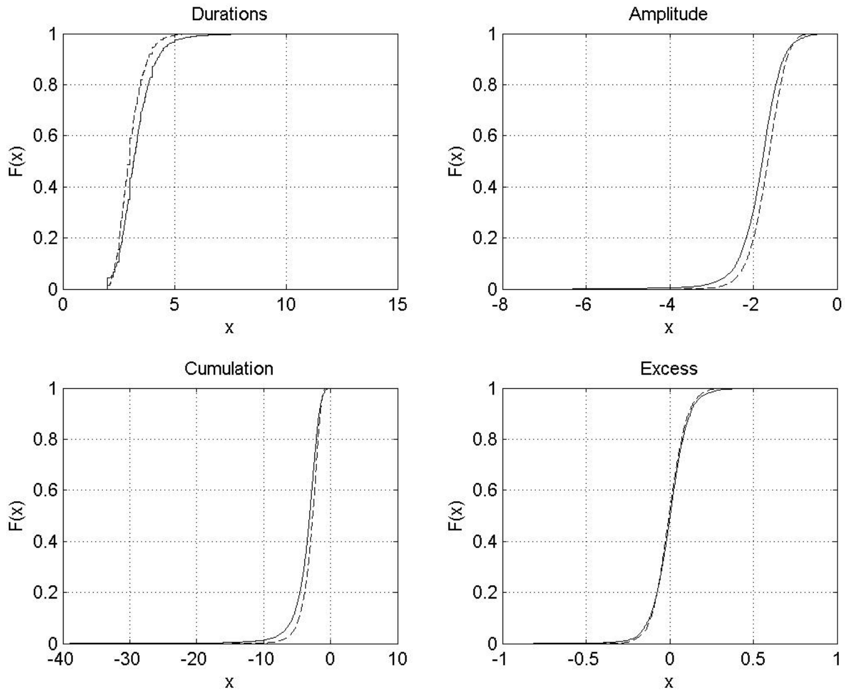


Figure A-5: U.K. Empirical Distributions of Average Recession Characteristics ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

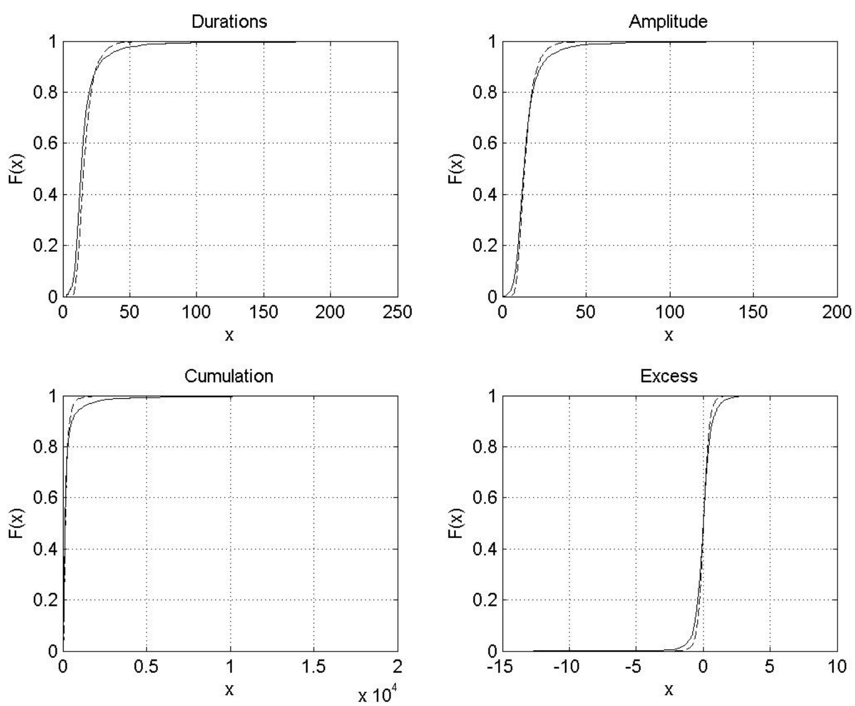


Figure A-6: U.K. Empirical Distributions of Average Expansion Characteristics ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

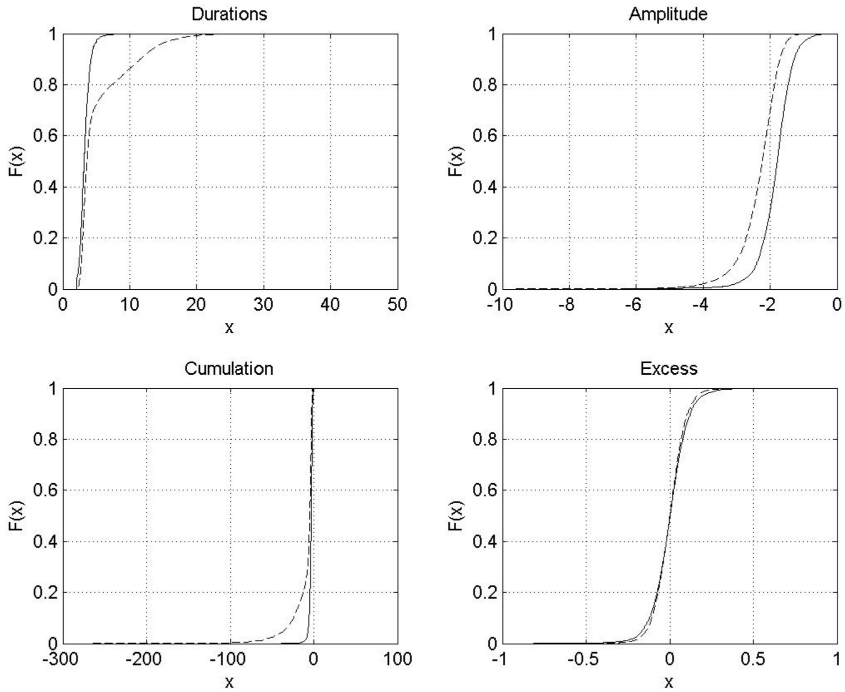


Figure A-7: U.K. Empirical Distributions of Average Recession Characteristics
MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

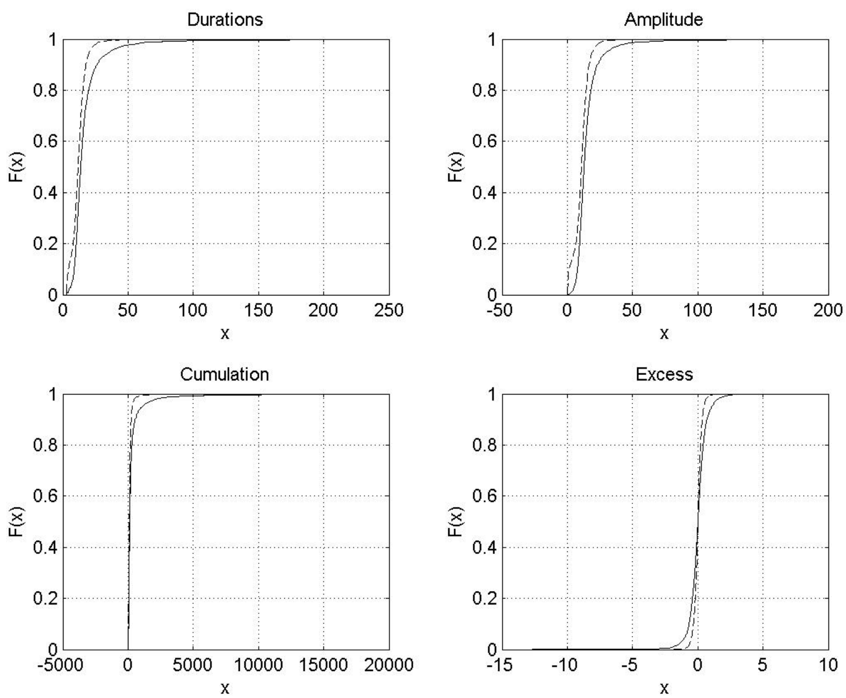


Figure A-8: U.K. Empirical Distributions of Average Expansion Characteristics
MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

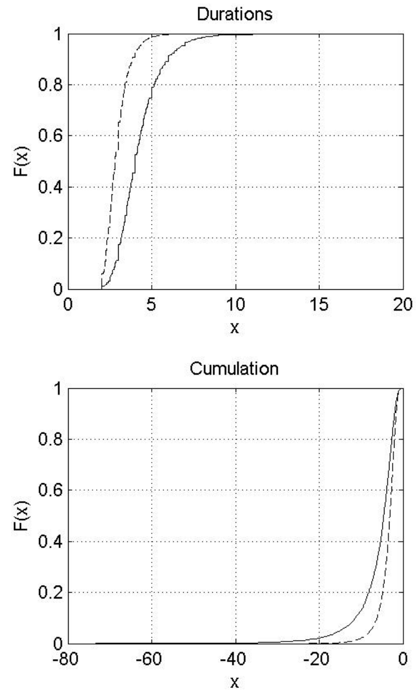


Figure A-9: Australia Empirical Distributions of Average Recession Characteristics ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

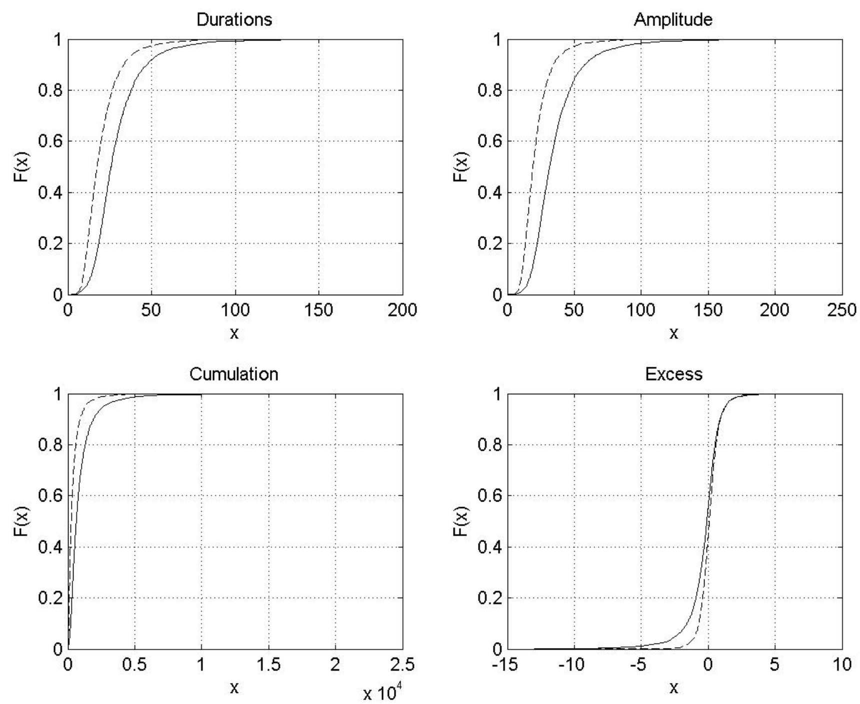


Figure A-10: Australia Empirical Distributions of Average Expansion Characteristics ARIMA-SV (dashed line) vs. MS-IS-NoAR (solid line)

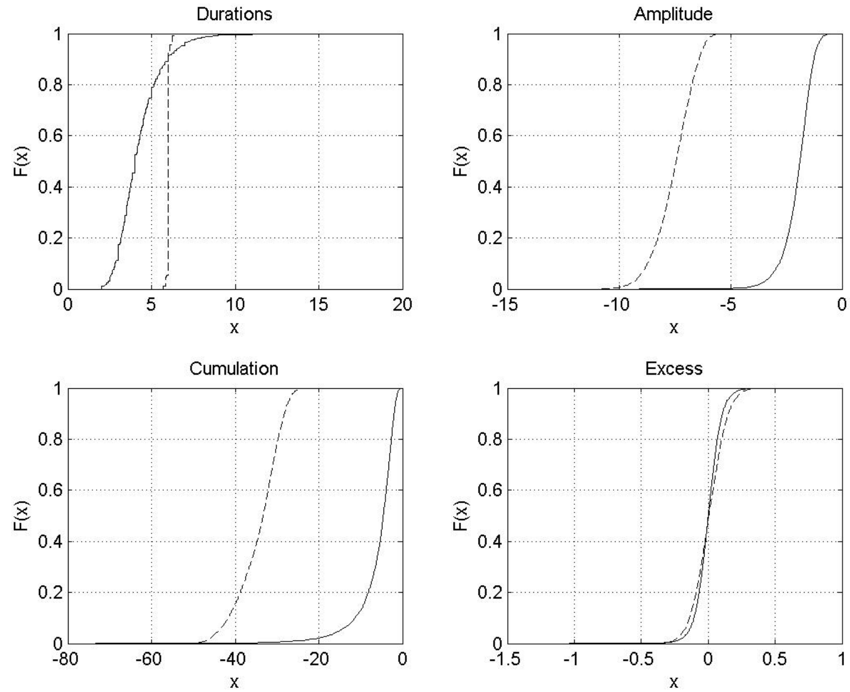


Figure A-11: Australia Empirical Distributions of Average Recession Characteristics MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

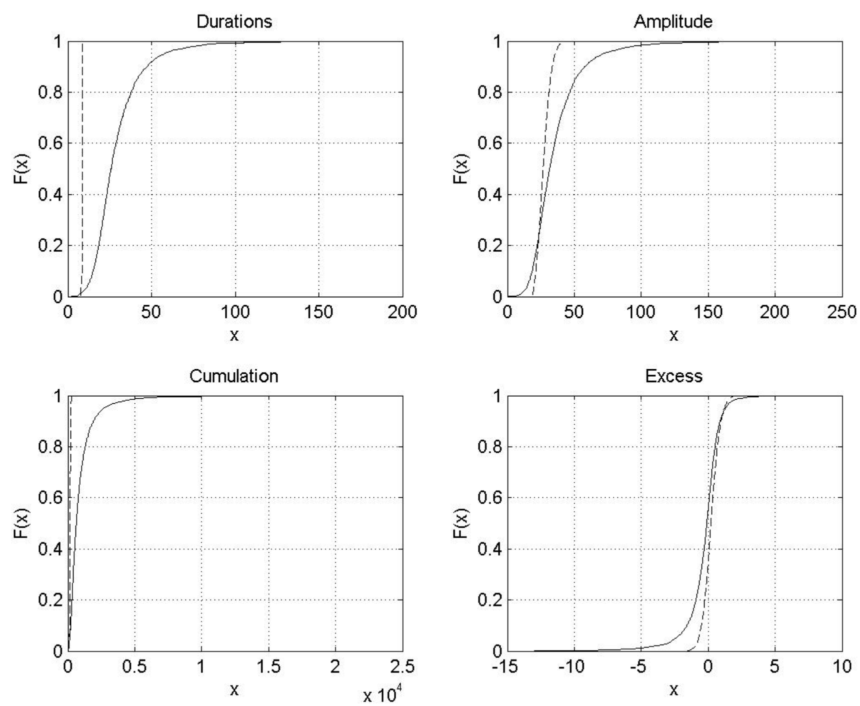


Figure A-12: Australia Empirical Distributions of Average Expansion Characteristics MS-IS-AR(1) (dashed line) vs. MS-IS-NoAR (solid line)

Bibliography

ADELMAN, I., AND F. L. ADELMAN (1959): "The Dynamic Properties of the Klein-Goldberger Model," *Econometrica*, 27(4).

ALBERT, J. H., AND S. CHIB (1993): "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics*, 11(1), 1–15.

ANDREWS, D. W. K. (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61(4), 821–856.

ANDREWS, D. W. K., AND W. PLOBERGER (1994): "Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), 1383–1414.

BRY, G., AND C. BOSCHAN (1971): *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. National Bureau of Economic Research, New York.

BURNS, A., AND W. MITCHELL (1946): *Measuring Business Cycles*. National Bureau of Economic Research, New York.

CAMACHO, M., AND G. PEREZ-QUIROS (2006): "A New Framework to Analyze Business Cycle Synchronization," in *Nonlinear Time Series Analysis of Business Cycles*, ed. by C. Milas, P. Rothman, and D. van Dijk, vol. 276 of *Contributions to Economic Analysis*. Elsevier, Amsterdam.

——— (2007): "Jump-and-Rest Effect of U.S. Business Cycles," *Studies in Nonlinear Dynamics and Econometrics*, 11(4), 1–37.

CANOVA, F. (1994): "Does Detrending Matter for the Determination of the Reference Cycle and the Selection of Turning Points?," Economics Working Papers 113, Department of Economics and Business, Universitat Pompeu Fabra.

——— (1998): "Detrending and Business Cycle Facts: A User's Guide," *Journal of Monetary Economics*, 41(3), 533 – 540.

Carrasco, M. (2002): "Misspecified Structural Change, Threshold, and Markov-switching models," *Journal of Econometrics*, 109(2), 239–273.

Cashin, P., and S. Ouliaris (2004): "Key Features of Australian Business Cycles," *Australian Economic Papers*, 43(1), 39–58.

Centre for Economic Policy Research (2009): "Euro Area Business Cycle Dating Committee," http://www.cepr.org/data/dating/info2.asp, CEPR Business Cycle Dating Committee.

Chauvet, M., and J. D. Hamilton (2005): "Dating Business Cycle Turning Points," Working Paper 11422, National Bureau of Economic Research.

Chauvet, M., and J. M. Piger (2003): "Identifying Business Cycle Turning Points in Real Time," *Review (00149187)*, 85(2), 47.

Clements, M. P., and H.-M. 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.

Cogley, T., and J. M. Nason (1995): "Output Dynamics in Real-Business-Cycle Models," *The American Economic Review*, 85(3), 492–511.

Davidian, M., and R. J. Carroll (1987): "Variance Function Estimation," *Journal of the American Statistical Association*, 82(400), 1079–1091.

Frisch, R. (1933): "Propagation Problems and Impulse Problems in Dynamic Economics," in *Economic Essays in Honor of Gustav Cassel*, pp. 171–205. Frank Cass and Company Ltd., London.

Galvão, A. B. C. (2002): "Can Non-Linear Time Series Models Generate US Business Cycle Asymmetric Shape?," *Economics Letters*, 77(2), 187–194.

Garcia, R. (1998): "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review*, 39(3), 763–788.

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

——— (2003): "Comment on A Comparison of Two Business Cycle Dating Methods," *Journal of Economic Dynamics and Control*, 27(9), 1691–1693.

——— (2005): "Regime-Switching Models," Discussion paper, University of California, San Diego, Prepared for Palgrave Dictionary of Economics available at dss.ucsd.edu/ jhamilto/palgrav1.pdf.

Hansen, B. E. (1992): "The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, S61–S82.

Hansen, B. E. (1997): "Approximate Asymptotic P Values for Structural-Change Tests," *Journal of Business and Economic Statistics*, 15(1), 60–67.

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

——— (2003a): "A Comparison of Two Business Cycle Dating Methods," *Journal of Economic Dynamics and Control*, 27(9), 1681.

——— (2003b): "Rejoinder to James Hamilton," *Journal of Economic Dynamics and Control*, 27(9), 1695.

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

Kim, C.-J., J. Morley, and J. Piger (2005): "Nonlinearity and the Permanent Effects of Recessions," *Journal of Applied Econometrics*, 20(2), 291–309.

Kim, C.-J., and C. R. Nelson (1999): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics*, 81(4), 608–616.

King, R. G., and C. I. Plosser (1994): "Real Business Cycles and the Test of the Adelmans," *Journal of Monetary Economics*, 33(2), 405–438.

Krolzig, H.-M., and J. Toro (2005): "Classical and modern business cycle measurement: The European case," *Spanish Economic Review*, 7(1), 1–21.

Leamer, E. E. (2008): "What's a Recession, Anyway?," Working Paper 14221, National Bureau of Economic Research.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?," *The American Economic Review*, 90(5), 1464–1476.

Morley, J., and J. Piger (2005): "The Importance of Nonlinearity in Reproducing Business Cycle Features," *The Federal Reserve Bank of St. Louis Working Paper Series*, Working Paper 2004-032B.

Murray, C. J. (2003): "Cyclical Properties of Baxter-King Filtered Time Series," *The Review of Economics and Statistics*, 85(2), 472–476.

National Bureau of Economic Research (2001): "The Business-Cycle Peak of March 2001," http://www.nber.org/cycles/november2001, NBER Business Cycle Dating Committee.

——— (2003): "The NBERs Recession Dating Procedure," http://www.nber.org/cycles/recessions.html, NBER Business Cycle Dating Committee.

——— (2009): "Business Cycle Expansions and Contractions," http://www.nber.org/cycles.html, NBER Business Cycle Dating Committee.

Neftci, S. N. (1993): "Statistical Analysis of Shapes in Macroeconomic Time Series: Is There a Business Cycle?," *Journal of Business and Economic Statistics*, 11(2), 215–224.

Sensier, M., and D. van Dijk (2004): "Testing for Volatility Changes in U.S. Macroeconomic Time Series," *Review of Economics and Statistics*, 86(3), 833–839.

Simkins, S. P. (1994): "Do Real Business Cycle Models Really Exhibit Business Cycle Behavior?," *Journal of Monetary Economics*, 33(2), 381–404.

Smith, P. A., and P. M. Summers (2002): "Regime Switches in GDP Growth and Volatility: Some International Evidence and Implications for Modelling Business Cycles," Melbourne Institute Working Paper Series WP2002n21, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

——— (2005): "How Well do Markov Switching Models Describe Actual Business Cycles? The Case of Synchronization," *Journal of Applied Econometrics*, 20(2), 253–274, Article.

van Dijk, D., D. R. Osborn, and M. Sensier (2002): "Changes in Variability of the Business Cycle in the G7 Countries," Discussion Paper 016, Economics, The Univeristy of Manchester, Centre for Growth and Business Cycle Research Discussion Paper Series.