ESSAYS ON THE DECOMPOSITION OF MACROECONOMIC TIME SERIES INTO PERMANENT AND TRANSITORY COMPONENTS

by

Christian Joseph Murray

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

1998

Approved by ______________________________
Chairperson of Supervisory Committee

Program Authorized
to Offer Degree Economics

Date June 12, 1998

UMI Number: 9907942

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature ____________

Date 6-12-98

University of Washington

Abstract

ESSAYS ON THE DECOMPOSITION OF MACROECONOMIC TIME SERIES INTO PERMANENT AND TRANSITORY COMPONENTS

by Christian Joseph Murray

Chairperson of the Supervisory Committee: Professor Charles Nelson
Department of Economics

This dissertation is comprised of three essays on modern macroeconometrics. The first essay takes issue with the conclusion of recent papers that U.S. output is trend stationary. It is shown that tests of the null hypothesis of a unit root against the alternative of trend stationarity are sensitive to data-based lag selection and departures from the maintained hypothesis of temporal homogeneity

Specifically, time series which contain a unit root may appear to be trend stationary if they are perturbed by large additive outliers. This generates the false appearance of trend stationarity. There is overwhelming evidence against the hypothesis of trend stationarity in the post-war data. Also, the implied business cycle is implausible.

The second essay extends the framework of dynamic factor models by specifying two common factors. Both common factors are subject to changes in regime. This allows for asymmetry in both the common permanent and common transitory components of time series. Previous work with dynamic factor models has restricted asymmetry to only the permanent component of a time series. This assumes *a priori* that recessions cause permanent damage. In light of recent evidence which suggests that recessions only temporarily lower output, we allow for the possibility that the

source of business cycle asymmetry is the common transitory factor. With the exception of the 1990-91 recession, which appears to be entirely permanent, we find that five of the six recessions from 1959 to the present are comprised of both permanent and transitory variation. Our parameter estimates imply that a six month recession permanently lowers the level of industrial production by 2.86%.

The third essay proposes a variety of statistics which test the hypothesis that a time series contains a point of structural change and/or has a unit root. Regarding tests for structural change, the literature is incomplete. A complete cataloging of these tests is undertaken. Also considered are statistics which test the joint hypothesis of a unit root and no structural change. While these have the potential to offer an increase in power over statistics which ignore the unit root hypothesis, this is generally not found to be the case.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank the members of my dissertation committee (Charles Nelson, Charles Engel, Richard Startz, Stephen Turnovsky, and Eric Zivot) and Chang-Jin Kim for their helpful comments on my research.

I am additionally indebted to Chang-Jin Kim and Eric Zivot. I spent many hours in their offices during which they were kind enough to teach me about statistics and modern time series analysis.

I would also like to acknowlege financial support from the Ford and Louisa Van Voorhis Endowment and the Grover and Creta Ensley Fellowship in Economic Policy at the University of Washington.

Finally, I am most grateful to my dissertation supervisor, Charles Nelson. Throughout all stages of my research, Charles was an superb mentor and a good friend. Indeed, working closely with Charles has been my most valuable experience in graduate school. As a result, all of my prior expectations of graduate school have been greatly exceeded. My only regret in graduating is that I will not be able to spend as much time with Charles.

DEDICATION

I wish to dedicate this dissertation to my loving wife Corrine.

INTRODUCTION

My training in economics at the University of Washington has essentially been an apprenticeship. That is to say I worked very closely with faculty on my research. Chapter 1 of this dissertation is based on joint research with my dissertation supervisor, Charles Nelson. This research began while I was Charles' research assistant during the second year of my graduate studies. Chapter 2 of this dissertation is based on joing research with Chang-Jin Kim. Chang-Jin is a former student from the University of Washington and is now at the Department of Economics at Korea University. Chapter 2 is a result of many conversations with Chang-Jin which took place while he was visiting the University of Washington during the third year of my graduate studies. Chapter 3 is based on joint research with Eric Zivot. Eric and I began working on trend break statistics during the second year of my graduate studies. Working closely with Charles, Chang-Jin, and Eric has been a tremendous pleasure that I shall always remember.

CHAPTER 1: THE UNCERTAIN TREND IN U.S. GDP

1.1 Introduction

Several recent papers have brought the literature full circle on the issue of whether the trend in U.S. real GDP is deterministic or stochastic. The modeling of aggregate output as transitory fluctuation around a deterministic trend was routine in empirical work until Nelson and Plosser (1982) showed that data for 1909-70 were consistent with the hypothesis that the trend is instead a non-stationary stochastic process akin to a random walk. Such processes contain a unit root in their autoregressive representation and require first differencing for stationarity. The model estimated by Nelson and Plosser implies that the stochastic trend contributes more to the variation in output than does the transitory component. They argued that an economic implication of this finding is that real shocks are much more important than previously thought, since it is presumably real shocks that impact the trend while monetary and fiscal shocks have only transitory effects.

Perron (1989) argued that by failing to allow for structural change, Nelson and Plosser vastly overstated the frequency of permanent shocks. He found that the same data reject the stochastic trend hypothesis in favor of the deterministic alternative if a break in the level of the trend is allowed to occur at 1929. His model implies that there has been one permanent shock to output during the 1909-70 period, that being a negative one, and that all other shocks have been transitory. Zivot and Andrews

(1992) showed that this finding still holds after critical values are adjusted to reflect data-based selection of the break date.

More recently, Ben-David and Papell (1995), Cheung and Chinn (1997), and Diebold and Senhadji (1996) have conducted tests with longer time series, extending U.S. output data back to 1870 and forward to the more recent past. All find that the longer time series strongly reject the stochastic trend hypothesis in favor of a deterministic trend without breaks. The implicit argument in these papers is that rejection of the unit root hypothesis can be attributed to an increase in power derived from a longer sample. These papers would thus suggest that as more data has become available, the evidence has become sharper, pointing now in the direction of determinism, leaving no role for permanent shocks.

Whether the trend in aggregate output is deterministic or stochastic has far-reaching implications for modeling the economy and for judging the success of macro-stabilization policy. The deterministic trend view implies that it is only because of transitory shocks, presumably primarily monetary and fiscal in origin, that the economy deviates from a smooth, constant-growth-rate path. The performance of monetary policy should then be measured by its success in achieving small departures from that path. If, on the other hand, shocks to the trend component are an important source of macro-economic fluctuations, then the modeling and identification of real shocks becomes critical for the conduct and evaluation of monetary policy. The two views of trend also have strikingly different implications for long run uncertainty: under the deterministic view, long run uncertainty is limited by the stationarity of the cycle, while under the stochastic trend view, uncertainty about future output grows without bound.

This paper examines the robustness of recent findings with respect to two issues: the finite sample implications of data-based model specification and the effect on test size of plausible departures from the maintained hypothesis that the data are generated by a homogeneous process. Section 1.2 of the paper reviews standard unit root tests on U.S. real GDP 1870-1994 and examines the data for homogeneity across sub-periods. Section 1.3 presents Monte Carlo experiments designed to study the two issues of size and robustness to departures from homogeneity. Section 1.4 focuses on the evidence from the post-war period which we regard as more likely to represent a homogeneous sample. Section 1.5 summarizes our results and presents our conclusions.

1.2. Trends and Non-homogeneity in U.S. real GDP

The evidence against the stochastic trend view is reflected in the test statistics shown in Table 1.1 for the annual U.S. real GDP series, 1870-1994, assembled by Maddison (1995). Before interpreting these results, we briefly review the tests and their maintained hypotheses.

Dickey (1976), Fuller (1976), and Dickey and Fuller (1979) developed a test of the null hypothesis that a unit root in the AR representation, rather than a deterministic trend, accounts for the non-stationarity of a trending time series. The Dickey-Fuller test runs the regression

$$y_t = \rho y_{t-1} + \alpha + \beta t + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t .$$

Under the unit root null, ρ=1, the first difference is a stationary AR process and the series is said to be "difference stationary" in the parlance of Nelson and Plosser. Under the alternative hypothesis ρ<1, the series is "trend stationary," a stationary AR process added to a deterministic linear trend. Dickey and Fuller showed that the t-statistic for testing ρ=1 has a non-standard distribution, and they tabulated Monte Carlo critical values for various sample sizes for a random walk with i.i.d. Normal shocks. They show that the limiting distribution remains the same when *k* lagged first differences "augment" the model to account for serial correlation (see also Hamilton (1994)).

In practice the lag length *k*, is unknown and is chosen by a data-dependent procedure. Building on work later published in Hall (1994), Campbell and Perron (1991) suggested starting with a maximum value of *k* chosen *a priori*, deleting lags until encountering a t-statistic indicating significance at the .10 level (greater than 1.645 in absolute value). This general-to-specific (GS) procedure has been followed by Perron (1989), Zivot and Andrews (1992), and others. Theoretical support for GS, as well as for various information criteria, was provided by Hall for the pure AR case and by Ng and Perron (1995) for the ARMA case. They showed that if the maximum lag allowed is at least as large as the true lag, then asymptotically inference is unaffected by the data-based lag selection.

Since we will be interested in departures from the maintained hypothesis of i.i.d. shocks in Dickey-Fuller, we also include the heteroskedastic-consistent unit root test of Phillips and Perron (1988). This test does not rely on a finite order AR representation, but instead employs a correction for serial correlation based in part on the spectral representation of the innovation sequence at frequency zero. The quadratic spectral kernel is used to estimate the spectrum, and Andrews' (1991) selection procedure is used to determine the number of autocovariance terms included in forming the estimate of the spectrum. The Phillips-Perron test has the same limiting distribution as the Dickey-Fuller test.

Perron (1989) provided a generalization of the Dickey-Fuller test to allow for the possibility of structural change taking the form of a one-time break in level, or slope, or both. In the case of a break in level only, which he considered appropriate for U.S. real GNP, the Perron test adds step and impulse dummy variables to the Dickey-Fuller regression:

$$y_t = \rho y_{t-1} + \alpha + \beta t + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \delta S(b)_t + \gamma I(b)_t + \varepsilon_t \, .$$

where S is zero through year b and one thereafter, and I is one in year $b+1$ only and zero otherwise. Under the unit root hypothesis, the impulse dummy accounts for a break in level, while under trend stationarity alternative, the step dummy does. Perron provided critical values under the maintained hypotheses that the break date is known, the innovations are i.i.d. Normal, and lag k is known.

The test of Zivot and Andrews (1992) differs from the Perron test in two regards. First, the null hypothesis is that the series has a unit root *and* does not contain a break; accordingly, their test regression does not include an impulse dummy. Second, Zivot and Andrews recognize that the break date is unknown *a priori* and estimate it to be that which maximizes the absolute value of the unit root test statistic. The test regression is

$$y_t = \rho y_{t-1} + \alpha + \beta t + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \delta S(\hat{b})_t + \varepsilon_t$$

where $\hat{b}$ is the estimated break date. Zivot and Andrews tabulate Monte Carlo critical values for $t(\rho=1)$ in the case of a random walk with i.i.d. Normal innovations and where k is assumed known to be zero. They confirmed Perron's choice of 1929 as the break date and his rejection of the unit root hypothesis for the Nelson-Plosser real GNP series. In practice, k is unknown and Zivot and Andrews did a GS search at each potential break date.

Leybourne and McCabe (1994) have developed a test of the null hypothesis that a series is trend stationary, with difference stationarity (a unit AR root) being the alternative hypothesis. They assume that the series has an unobserved components representation where the trend is a random walk, the stationary component is AR(k), and the innovations are independent across components and are i.i.d. This implies that the univariate representation of the first differences is ARMA(k,1) and the MA part will have a unit root if the trend is deterministic (zero variance in the random walk). Critical values for the Leybourne-McCabe test are tabulated in Kwiatowski et al.

(1992). As in the Dickey-Fuller test, the Leybourne-McCabe test necessitates the preliminary step of selecting the lag length k to account for serial correlation. However, under the null hypothesis the series follows a non-invertible ARMA(k,1) process. The distribution of the AR terms is thus unknown. Therefore, in contrast to the Dickey-Fuller test, there does not exist a set of results which guarantees that inference based on the Leybourne-McCabe test is asymptotically unaffected by data-based lag selection.

Results of these tests are reported in Table 1.1 for the full Maddison sample and the sub-period 1909-1970 studied by Nelson and Plosser. The lag length is chosen alternatively by GS and Schwarz' (1978) information criteria (SIC). Following Perron and Zivot and Andrews, the maximum lag we consider for annual data is 8. As noted before, if the true lag is less than or equal to 8, the results of Hall (1994) and Ng and Perron (1995) state that in the limit, both GS and SIC will choose the correct lag with probability one. The Zivot-Andrews procedure identifies 1929 as the break date for both time spans. P-values are obtained by simulation as in the original articles; the DGP under the null hypothesis is a random walk for the unit root tests and a trend plus random error for Leybourne-McCabe.

Several features of the results seem worthy of note:

1. The null hypothesis of a unit root is strongly rejected by all three tests for the Maddison data. The Dickey-Fuller test for the sub-period studied by Nelson and Plosser is less favorable to the unit root hypothesis than they reported using data available prior to the work of Balke and Gordon.

2. Evidence against a unit root is stronger for the full time period.

3. Rejection of the unit root null is stronger when a break is allowed if the break date is assumed known.

4. The choice of lag k differs greatly between GS and SIC selection, the former often choosing the maximum allowed while the latter in every case chooses only one lag.

5. Lag selection matters for inference. For the Perron and Zivot-Andrews tests, GS lag selection leads to stronger rejection of the unit root. For the Leybourne-McCabe test, lag selection determines the outcome. SIC does not lead to rejection of trend stationarity (as reported by Cheung and Chinn, 1997), but GS does.

6. The step dummy is highly significant by conventional standards in every case However, Banerjee, Lumsdaine, and Stock (1992) demonstrate that the distributions of break-dummy coefficients are non-standard.

These tests have as maintained hypothesis that the series is homogeneous, generated by an AR process of known order with constant parameters and i.i.d. Normal innovations. It is not clear how deviations from these maintained hypotheses might affect size or power, although recent contributions to the theoretical literature, discussed below, suggest that they will. This is a concern in the context of U.S. GDP since 1929, when methods of data collection change, the Great Depression, and World War II are points at which the GDP data process might be expected to exhibit changes in both volatility and serial correlation.

For the period to 1929, Maddison used estimates by Balke and Gordon (1989); an alternative series is by Romer (1989). Both build on the pioneering methodology of Kuznets (1941,1946) and extensions by Kendrick (1961) and Gallman (1966). Briefly, the Kuznets methodology relies on trends extrapolated between benchmark years, then deviations from trend are based on indicator variables such as commodity output. It would be surprising if this method did not affect serial correlation and unit root tests. Indeed, the Dickey-Fuller test applied to the Balke-Gordon data rejects the unit root at the 10% level.

The period immediately following 1929 was one of banking failures on an unprecedented scale and repeated failure of the new Federal Reserve System to stabilize the system (Friedman and Schwartz, 1963). By the time the economy had recovered from the Great Depression it was jolted by World War II. The magnitude of these shocks is apparent in Figures 1.1 and 1.2 for levels and growth rates, and in the summary statistics in Table 1.2 for growth rates and deviations from the least squares trend line. The largest observations during the Depression-WWII period are three to four times the sample standard deviation measured over the full period. Fitted AR(3) models reveal large changes in serial correlation and much higher residual variance during the 1930-45 sub-period. The extremely small probability of this occurring in a homogeneous Normal sample is reflected in the asymptotic Jarque-Bera p-values of zero for the full sample. The three fold increase in the standard deviation of shocks is comparable to that for stock returns reported by Schwert (1989b). Separating pre-1929, 1930-45, and post-war periods, however, one obtains apparently homogeneous Normal samples.

More formal evidence on non-homogeneity comes from an extension of the Clark (1987) model in which stochastic trend and cycle components are augmented by an additive irregular component that switches on and off according to a Markov process. Details are given in Appendix A and in Murray (1997). As seen in Figure 1.3, the irregular component switches on in 1893, but shows only small fluctuations until 1930 when it reflects the huge swings in output of the Depression and WWII. Then it switches off in 1947. This pattern is consistent with larger measurement errors in the pre-1929 data and the 1930-46 period characterized by a sequence of additive outliers that do not occur elsewhere. Further, the cyclical component is no longer significant once this irregular component is included in the model. Our results are consistent with those of Balke and Fomby (1991) who also identified additive outliers associated with the Depression and World War II, but failed to detect permanent breaks in level.

1.3 The Sensitivity of Unit Root and Trend Stationarity Tests to Lag Selection and Additive Outliers

1.3.1 Design of the Experiments

This section presents a series of experiments designed to investigate how tests for a unit root or trend stationarity are affected by data-based lag selection and departure from the i.i.d. Normal assumption in the form of additive outliers. Our strategy is to specify a data generating process for 1870-1994 that contains a unit root and replicates the main statistical features of post-war GDP and use it to study size or power under GS and SIC lag selection, then introduce various types of additive outliers to see what effect they have on the tests. Our choice of the post-war data as a guide to the DGP is based on the results of the Markov-switching state-space model discussed above. We do not attempt to model the measurement errors in the pre-1929 data.

For the post-war Maddison data in first differences of logs, SIC chooses ARMA(0,0) and AIC chooses ARMA(1,0). We adopt the AR(1) model to include some dynamics. The estimated model, and our underlying DGP, is:

$$\Delta y_t = .17\Delta y_{t-1} + .027 + \varepsilon_t$$

$$\varepsilon_t \sim \text{i.i.d. N}(0, 0.025)$$

This DGP is run for 20 periods before recording a realization corresponding to 1870-1994. After integrating the first differences to obtain levels, we add one of a number of types of outliers to see its effect on the test statistics. The observed data, say y*, is then

$$y_t^* = y_t + O_t$$

where the outlier sequence $\{O_t\}$ varies across experiments. Two panels give results for sample lengths of 125 and 62 years corresponding to the full 1870-1994 period and the Nelson-Plosser 1909-1970 sub-period, respectively, based on 1,000 replications. The upper bound for the standard error of the rejection frequencies we report is .016 (see Davidson and MacKinnon (1993)).

In the first experiment, reported in Table 1.3, we have subtracted a fixed quantity from the level of simulated log real GDP in 1930 only. The value of O_{1930} ranges in successive experiments between 0 and -.4, the latter representing a one third reduction in output. The dip lasts for only one year, so the observed series resumes its underlying path in 1931 with no permanent change in level. In the second experiment, reported in Table 1.4, we consider outlier events that begin in 1930 and then follow either fixed or stochastic paths.

Reported in each table are actual frequencies of rejection of the unit root hypothesis at a nominal .05 significance level based on critical values reported in Fuller (1976), Zivot and Andrews, and Perron (with corrections to the Perron critical values by Zivot and Andrews), respectively. For the Perron and Zivot-Andrews tests the critical values are asymptotic, but for the Dickey-Fuller test the finite sample critical values are exact if k is known to be zero and the innovations are Gaussian. We also report the frequency with which the t-statistic for the step dummy in the Perron and Zivot-Andrews regressions is significant at the conventional nominal .05 level.

In the case of the LM test, the null hypothesis is trend stationarity, so the frequency of rejection is the power of the test against the alternative represented by our DGP. Since these frequencies are meaningless unless the size of the test is correct, we follow Cheung and Chinn in setting critical values by simulation of the trend stationary AR model suggested by the historical data.

The number of lagged first differences included in any of these regressions, denoted k, is selected alternatively by GS and SIC procedures described above. Note that in searching for the break date in the Zivot-Andrews test, the selection of k is repeated at every potential break date.

1.3.2. The Effect of Data-based Lag Selection on Test Size or Power

In the experiment reported in the first line of each panel of Table 1.3, no outlier has been added. In this case the series does in fact have a unit root with i.i.d. Normal innovations, so the frequency reported for each unit root test is its actual size, the probability of rejecting the null hypothesis at a nominal .05 level when it is true. As reported previously by Hall (1994) for the Dickey-Fuller test, size depends importantly on the method of lag selection. In the case of the Dickey-Fuller and Perron tests, SIC produces roughly the correct size, while GS results in a size of about .10. The Zivot-Andrews test suffers from greater size distortion under both lag selection strategies, and the distortion is entirely due to selecting k from the data; when the correct value $k=1$ is imposed, we find that the actual size is correct.

While Hall showed that both GS and SIC are valid asymptotically, his Monte Carlo results demonstrated that there may be substantial size distortions in finite samples such as we see here. It is clear that lag selection is not a simply a trade-off between size and power, with strategies favoring large k offering more correct size but lower power. The analogy to including extraneous variables in a regression which use up degrees of freedom but do not create a bias is misleading because the *particular* value of k is based on pretesting. If k were set *a priori* and was larger than the true value of k in any particular case, then the test *would* have correct size but lower power; see Ng and Perron (1995). An appropriate analogy is to the problem of data-based selection of instruments in 2 stage least squares, where Hall, Rudebusch, and Wilcox (1994) have shown that searching for the best instruments severely distorts the size of tests on structural coefficients.

The Phillips-Perron test relies on the data to select lag length for truncation of the autocovariance function used in estimation of the spectrum at frequency zero, rather than for selecting AR order. It is the only one of the tests considered here that has too small a size. The size distortion of this test is evidently dependent on the form of the autocorrelation function, since Schwert (1989a) found that the size of this test was too large in the MA(1) case.

As noted above, Banerjee, Lumsdaine, and Stock (1992) demonstrate that the asymptotic distributions of break dummy coefficients are non-standard, although they maximized the F-stat of the dummy variable, rather than the unit root statistic, across break dates. Using the t-distribution for inference would create the false impression that a break occurred. As seen here, the sizes of the t-tests for the step dummy are much too large; about 0.2 in the Perron regressions and about 0.95 in the Zivot-Andrews regressions, sample length seeming to have little influence. The size of the t-test for the impulse dummy in the Perron regression is also excessive (but not shown).

In the case of the Leybourne-McCabe test, the frequency of rejection reflects power against the alternative of the ARIMA(1,1,0) process generating the data. We note that power is higher under SIC lag selection. For T=125, the test correctly rejects trend stationarity about 80% of the time using SIC and about 60% for GS. For T=62, the power is significantly decreased; 44% for SIC and 24% for GS.

1.3.3 The Effect of Additive Outliers

Successive experiments reported in Table 1.3 add -0.1,-0.2,-0.3, or -0.4 to the level of simulated log real GDP in 1930 only, after which it returns to the underlying process. This range of outliers was motivated by the range of extreme values reported in Table 1.2 for the 1930-1945 period. The resulting process still contains a unit root, but it is no longer homogeneous.

As the magnitude of the outlier increases, rejection rates for all of the unit root tests rise sharply. Frances and Haldrup (1994) studied the effect of a stochastic additive outlier that occurs with some probability each time period in an *I(1)* process, and showed that even asymptotically the distribution of the Dickey-Fuller t-statistic (k fixed) is shifted to the left, increasing the frequency of rejection of the unit root null. A stochastic outlier introduces an MA(1) component into the process, so no finite AR representation exists and this is the situation in which Schwert (1989a) had shown that unit root tests have poor finite sample properties. Whether the outlier is stochastic or fixed, the maintained hypothesis of a finite order AR with i.i.d. Normal errors is violated, and rejection is triggered.

It is completely general that rejection of a null hypothesis does not imply that the alternative hypothesis depicted by the test regression is true. Faced with the choice between the unit root null and the trend stationary alternative when neither is true, these tests reject the unit root. If misinterpreted, these tests spuriously signal trend stationarity, with a level break if allowed, when in fact the series has a unit root and the outlier event affects only one observation. It is interesting that frequencies of rejection of the unit root diminish when the sample size is doubled, seemingly at odds with the idea that the power of a test should increase with sample size. However, the present case is one where the null hypothesis also becomes less wrong as sample size grows, since the departure from homogeneity becomes relatively less severe the longer the time series.

For the Leybourne-McCabe test, the issue is how power is altered by introduction of the outlier. For the one period outliers considered in this section, the power is marginally reduced for T=125, while T=62 results in a more severe power reduction.

1.3.4 Additive Outliers That Persist

The experiments reported in Table 1.4 are designed to explore the impact of outlier duration and pattern. Starting from the benchmark case of no outlier, we add -.2 in 1930 only (as in Table 3), in each of the ten years 1930-39, and to every year from 1930 on. The last case is motivated by Perron's (1989) finding of a permanent break in the level of output in 1930. Further experiments adds a stochastic additive outlier sequence generated by an AR(2) modeled on departures from a local trend connecting 1929 to 1946 or, alternatively, the fixed, actual de-meaned cumulative changes from 1930 through 1945. This last case is in the spirit of the experiments reported by Kilian and Ohanian (1996).

Note that persistence in itself reduces the frequency of rejection of the unit root in the Dickey-Fuller, Phillips-Perron, and Perron tests. When the outlier is permanent, the size of the Perron test is close to the correct 0.05. Indeed, the Perron regression is correctly specified in that case with inclusion of the impulse dummy at the correct date, allowing for a permanent change in level. In contrast, the Zivot-Andrews test rejects the unit root null much more frequently if the level of the series shifts permanently, and the step dummy is almost always significant. It turns out that this is not due to the absence of an impulse dummy in the Zivot-Andrews regression; when it was rerun with the impulse dummy the results were essentially the same. Rather, it is having to search for the break date (which is identified quite poorly) that accounts for excessively frequent rejection of the unit root. This suggests a high value for information about the timing of structural change.

In the last two experiments in Table 1.4, the underlying process is inundated with a high amplitude wave which distorts the level of the series for 16 years and then is gone. In both, the unit root and no-step-dummy null hypotheses are rejected often, except by Phillips-Perron. Rejection rates differ sharply depending on which lag selection method is used. GS generally chooses a much larger value of k than does SIC, reflecting the contrast we saw in Table 1.1. It also appears that the *particular* pattern of real GDP during the period 1930-45 as opposed to the random outcomes of the AR(2) process do matter; the tendency to stronger rejection of the unit root in the longer sample being more apparent for the fixed pattern.

Finally, the Leybourne-McCabe test has substantially lower power in the last two experiments, failing to reject trend stationarity much more frequently if the underlying unit root process is overlaid by a transitory component of large amplitude. This is not surprising, since the stochastic trend will appear relatively smooth compared to the transitory component. As pointed out by Cochrane (1991), there is an observational equivalence between a trend stationary process and one with a stochastic trend where the variance of the innovations is small enough relative to the transitory component.

1.4 Evidence from the Post-war Data

The argument that a longer span of data yields a test statistic with greater power is valid only if a time series is temporally homogeneous. Thus while more data is usually preferred to less, we find two compelling reasons to focus on post-war GDP data for testing the unit root hypothesis. Recall that the Balke-Gordon data used by Maddison up to 1929 are constructed by linear interpolation between benchmark years, so unit root tests may be biased toward rejection. During the next 16 years the economy was subject to the large shocks associated with the Great Depression and World War II. The experiments reported above imply that even if these events were entirely transitory, they could account for rejection of the unit root hypothesis and be misconstrued as evidence of trend stationarity with or without structural change. Thus by focusing on the post-war data, we hope to minimize the chance of spuriously rejecting the unit root hypothesis due to violation of the homogeneity assumption.

To study the post-war period we shift from the annual Maddison data to the recently available quarterly series of real GDP in chained (1992) dollars. One advantage of the chained data for our purposes is that it addresses the concern of Gordon (1993) that a productivity slow-down would be obscured in a series based on fixed weight price deflators, such as the real GDP data used by Maddison. Indeed, the existence and causes of such a structural break have been discussed since the 1970s and are the subject of a large and continuing literature; see also Baily and Gordon (1988). Indeed, Perron reported evidence of a break in 1973 in the *slope* of the trend function for post-war quarterly real GNP, indicating a slow down in long-term growth, although Zivot and Andrews were not able to confirm this finding when they searched for the break date. Results of unit root, trend break, and trend stationarity tests for chained post-war quarterly real GDP are presented in Table 1.5. Nominal p-values are based on tabulated asymptotic distributions, while exact p-values will depend on the method of lag selection and are obtained by simulation under the null hypothesis where the data generating process is either an AR(1) model fitted to first differences of the log of the post-war data, or an AR(2) around a deterministic time trend, the orders chosen both by GS and SIC.

1.4.1 Unit Root Tests and Confidence Intervals for the Largest AR Root

The Dickey-Fuller tests reported in Table 1.5 are fully consistent with a unit root in post-war real GDP. Following Perron and Zivot and Andrews, we start with a maximum lag of 12 for quarterly data. For both methods of lag selection the nominal and exact p-values are greater than 0.50; the expected value of the test statistic being about -2.2. For completeness, we include the Phillips-Perron test, although its nominal size is too small, and it too is fully consistent with a unit root. A common criticism of Dickey-Fuller tests is that they have low power against local alternatives, an AR root close to unity. A modified test by Elliot, Rothenberg, and Stock (1996), which they call DF-GLS$^{\tau}$, employs a local-to-unity detrending procedure designed to maximize power against local alternatives. Although lag selection differs sharply between GS and SIC, the test results do not, both being entirely consistent with a unit root.

Although these results are consistent with a unit root process, they are also consistent with a range of trend stationary alternatives since it is not possible to distinguish in a finite sample between the realization of a unit root process and a trend stationary process with an AR root close enough to unity. This is the observational equivalence problem identified by Nelson and Plosser and emphasized by Christiano and Eichenbaum (1990) among others. As noted above, Cochrane (1991) has identified a corresponding observational equivalence between a trend stationary process and one with a stochastic trend where the variance of the innovations is small enough, and Engel (1997) shows that an economically significant random walk component can be missed. Thus, the range of models that cannot be rejected by any finite data set must always include both unit root and trend stationary alternatives. We would like to know how wide that range is in any given case.

To see the range of the largest AR root, ρ, that is consistent the post-war chained GDP data, we computed two-sided confidence intervals using the procedure developed by Stock (1991). These are based on inverting the augmented Dickey-Fuller test statistic to determine the values of ρ consistent with it. The 95% interval based on GS selection of 12 lags is (0.961, 1.026) and based on SIC selection of 1 lag it is (0.931, 1.022). We note that both include trend stationary as well as explosive alternatives. The former possibility has received considerable attention in the recent literature (Rudebusch 1992, 1993). Rudebusch (1993) demonstrates that the augmented Dickey-Fuller test applied to post-war quarterly GNP lacks power against a specific non-local alternative. He fits an AR(2) to deviations from the linear trend and shows that for this parameterization a unit root statistic greater than or equal to the observed statistic occurs 22% of the time, suggesting that the distribution of the statistic is not radically different under the unit root and this stationary alternative. Applying Rudebusch's technique to the chained data, we find that the trend stationary representation yields statistics greater than or equal to the observed unit root statistics 2.47% and 14.55% of the time for GS and SIC respectively. Thus, the disparity between the unit root and trend stationary alternative is greater in the chained data.

1.4.2. Implications of Trend Stationarity for the Cyclical Behavior of GDP

While Jones (1995) and Diebold and Senhadji (1996) argue for the efficacy of trend stationarity in forecasting the long run path of output, little attention has been given to the particular realization of the transitory component that is implied for the post-war U.S. and whether it corresponds to an economically meaningful deviation from a long run growth path. Figure 1.4 plots the deviation of the log of chained GDP from the fitted trend line, with NBER reference cycles shaded. While the deviation from trend does dip in concert with NBER recessions, its variation is dominated by a very low frequency wave that says that the economy was well below trend most of the period from 1947 through the early 1960s, consistently above trend until 1981, finally falling sharply below trend during the last recession and continuing downward through 1997. Conventional measures of economic performance would suggest a very different pattern, unemployment having been very high during the 1974-75 and 1981-82 recessions and very low in 1997. The implied deviations from trend are also of large amplitude, starting at -7% in 1947, peaking at +10% in 1973, and ending at -8%

in 1997. We are not aware of any estimates that an additional 8% of output was available to the U.S. economy in 1997. A forecast based on trend stationarity would imply growth rates about one percentage point above average for the next several years.

The impression of a long wave in Figure 1.4 is reflected in the long wave observed in the correlogram in Figure 1.5 and the low frequency peak in sample spectrum plotted in Figure 1.6. These features are reminiscent of the spurious periodicity, identified by Nelson and Kang (1981), that characterizes residuals from the regression of a random walk, and *I(1)* processes in general, on time. They show that the spurious cycle typically has a period equal to about 0.83 of the length of the series, here about 42 years. Indeed, the peak of the sample spectrum occurs at a frequency of 0.035, which implies a period of 45 years, slightly above that predicted by Nelson and Kang. These low frequency dynamics, as well as the economic implausibility of the implied cycle, suggest to us that the trend component of output is much more flexible than a straight line, probably accounting for much of the long wave that trend stationarity would attribute to the transitory component we see in Figure 1.4.

1.4.3 How Big is the Random Walk in GDP?

Cochrane (1988) criticized the use of unit root tests to determine the long run dynamic properties of a time series. Since unit root tests rely on parsimonious representations of the short run dynamics, they only use the first few terms of the autocorrelation function and may fail to capture the long run behavior of a time series. Cochrane advocated a non-parametric measure of long run persistence, the ratio of the variance of the j^{th} difference to the variance of the first difference, normalized by the factor $1/j$. If a series is trend stationary, the variance ratio approaches zero as $j \rightarrow \infty$. If a series is integrated, it can be decomposed into a random walk plus a stationary component (Beveridge and Nelson (1981)) and the variance ratio then approaches the ratio of the variance of the random walk to the variance of the first difference, so it is unity for a pure random walk. Thus, the variance ratio provides an estimate of the contribution of the stochastic trend to the long run dynamics of a time series. The sample variance ratio using Cochrane's unbiased estimate (his equation A3) for the post-war chained GDP is plotted in Figure 1.7 and, unlike the shorter series used by Cochrane, shows no tendency to decline at longer lags, suggesting that the variation in GDP is dominated by the variation in the stochastic trend.

1.4.4 Is There a Productivity Slow-Down in Chained GDP?

We now turn to the issue of a productivity slow-down in the U.S. economy and any implications it might have for tests of the unit root hypothesis. It is a fact that growth has been slower since 1973: the annual growth rate over the period 1947.1-1973.1 was 3.9% while in 1973.2-1997.3 it fell to 2.5%. Whether this difference is statistically significant and, if so, whether it represents an abrupt structural change or a gradual evolution toward slower growth is unclear. Model B of Perron allows for a break in the growth rate under the trend stationary alternative, though not under the null. It differs from Model A in replacing the step and impulse dummies with a "ramp" dummy that is zero through the break date then increasing arithmetically, so the trend function is allowed to bend but not shift. Perron applied this test to post-war quarterly real GNP, 1947-86, setting the break date at 1973:1, and rejected the null hypothesis. Zivot and Andrews estimated the break date at 1972:2 but did not reject the unit root. Both used GS, choosing $k = 10$ quarters.

For the post-war chained data, both GS and SIC provide no evidence against the unit root hypothesis. As seen in Table 1.5, 1972.2 is chosen as the break date as in Zivot and Andrews, and the nominal p-values are 0.54 and 0.14 respectively. This corroborates the finding of Zivot and Andrews that post-war GDP is not well characterized as stationary fluctuations around a kinked time trend. Exact p-values reflect the finite sample size distortion induced by lag selection.

1.4.5 The Leybourne-McCabe Test for Trend Stationarity

Finally, Table 1.5 also reports the results of the Leybourne-McCabe (1994) test for trend stationary applied to the post-war chained data. As noted by Cheung and Chinn (1997), the asymptotic p-values provided by Kwiatowski et. al. (1992) are not useful guides for inference when the sample is finite. Indeed, the Leybourne-McCabe statistics based on GS and SIC lead to rejection of the trend stationary null at any significance level. We also computed exact p-values for the observed statistics based on the trend stationary AR(2) parameterization discussed above. The likelihood of observing the Leybourne-McCabe statistics under GS and SIC is 9.3% and 2.1% respectively, offering little evidence in favor of pure trend stationarity for the post-war data.

1.5 Conclusion

Recent research has demonstrated that standard tests reject the null hypothesis of a unit root in U.S. real GDP over the period 1870-1994 in favor of the alternative of stationarity around a log-linear trend. If valid, these findings would imply that all shocks are temporary and that the long run path of the economy is deterministic. This paper calls that inference into question on two grounds.

First, the size of these tests is distorted in finite samples by the necessary preliminary step of selecting the number of lagged first differences to be included in the regression. We show that, for parameterizations suggested by the data, the actual probability of rejecting the unit root hypothesis when it is true is substantially greater under data-based lag selection than is indicated by the nominal significance levels upon which rejections of the unit root have been based in the recent empirical literature.

Second, the long historical time series used in the literature violate the maintained hypothesis that the data generating process is temporally homogeneous. The period 1930-45 was one of unusually large disturbances that may have been largely temporary in their effect on the level of output. However, we find that outliers added to the level of a unit root process for only one period are sufficient to trigger rejections of the unit root hypothesis with high probability. Given the choice between two wrong models, the unit root tests lean towards trend stationarity although it is false.

To reduce the possibility of spurious rejection of the unit root null due to heterogeneity in the data, we focus on post-war chained GDP. The unit root statistics in all cases not only fail to reject, but lie in the upper half of the distribution under the null hypothesis. While we also cannot reject a range of trend stationarity alternatives, we find that the implied cycle component contains a low frequency peak in the sample spectrum with a period of 45 years, much longer than the 6.5 year average peak-to-peak length in the NBER chronology. This is reminiscent of the spurious periodicity phenomenon analyzed by Nelson and Kang (1981) for a detrended unit root process. Furthermore, the cycle implied by detrending post-war GDP contradicts employment based measures of economic activity; it implies below-trend performance during the 1960s, above-trend performance in the 1970s, and then a decline that puts real GDP 8% below trend in 1997. These results cast serious doubt on the trend stationary model as an economically credible representation of real GDP.

In our view, a constructive direction for modeling aggregate output will be one that moves beyond the unit root issue and the use of dummy variables to represent shifts in level or growth rate. Determinism is not an hypothesis that is supported either in economic theory or in history. Dummy variables restrict the frequency of permanent shocks, and give no guidance as to the likelihood or size of future shocks. A statistical model implies a conditional distribution of future observations given the data, not simply an accounting of past events.

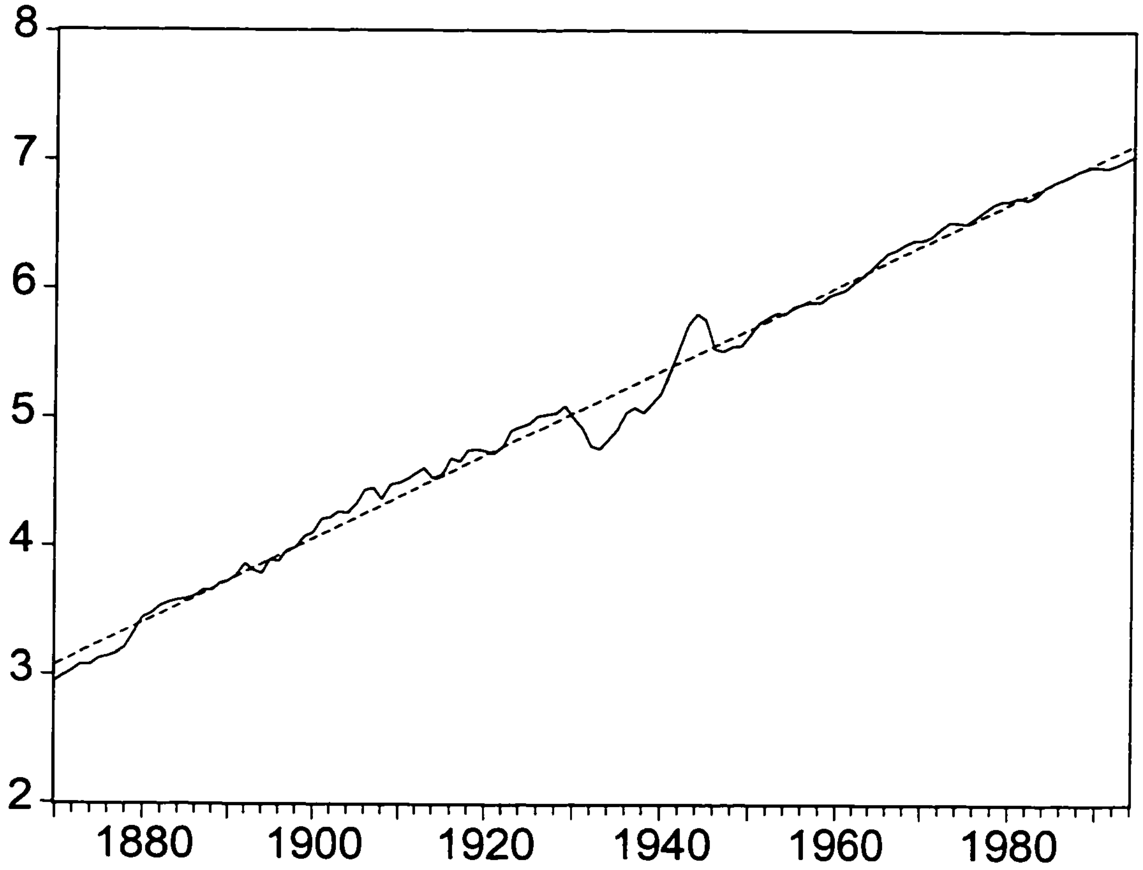

Figure 1.1 Log of Real GDP; Maddison (1995) Data

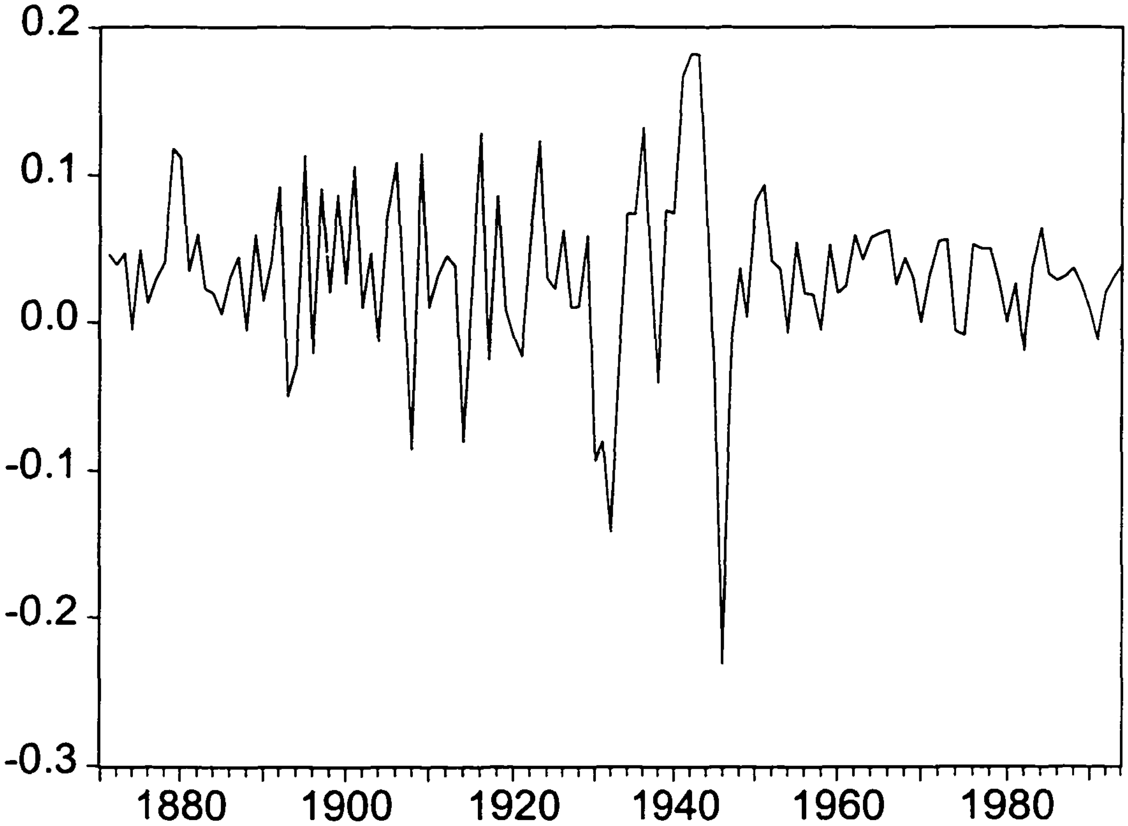

Figure 1.2 Growth Rate of U.S. Real GDP

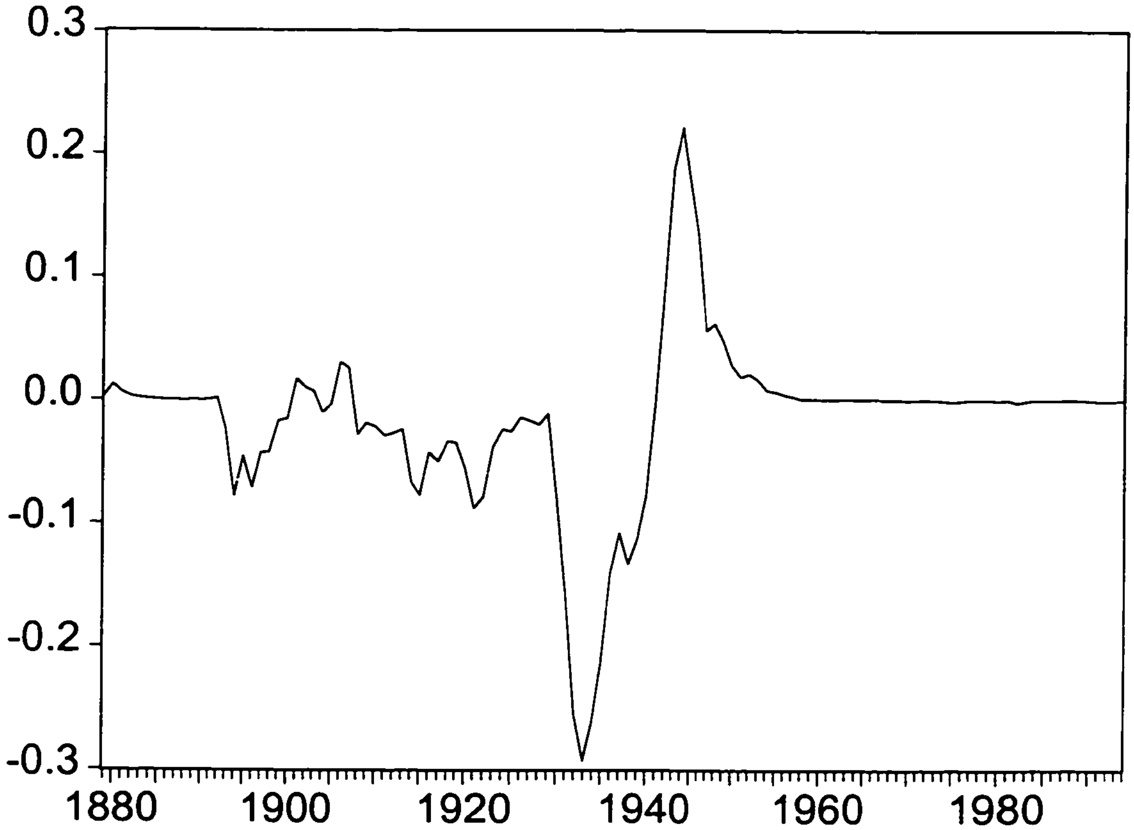

Figure 1.3 Irregular Component, U.S. Real GDP

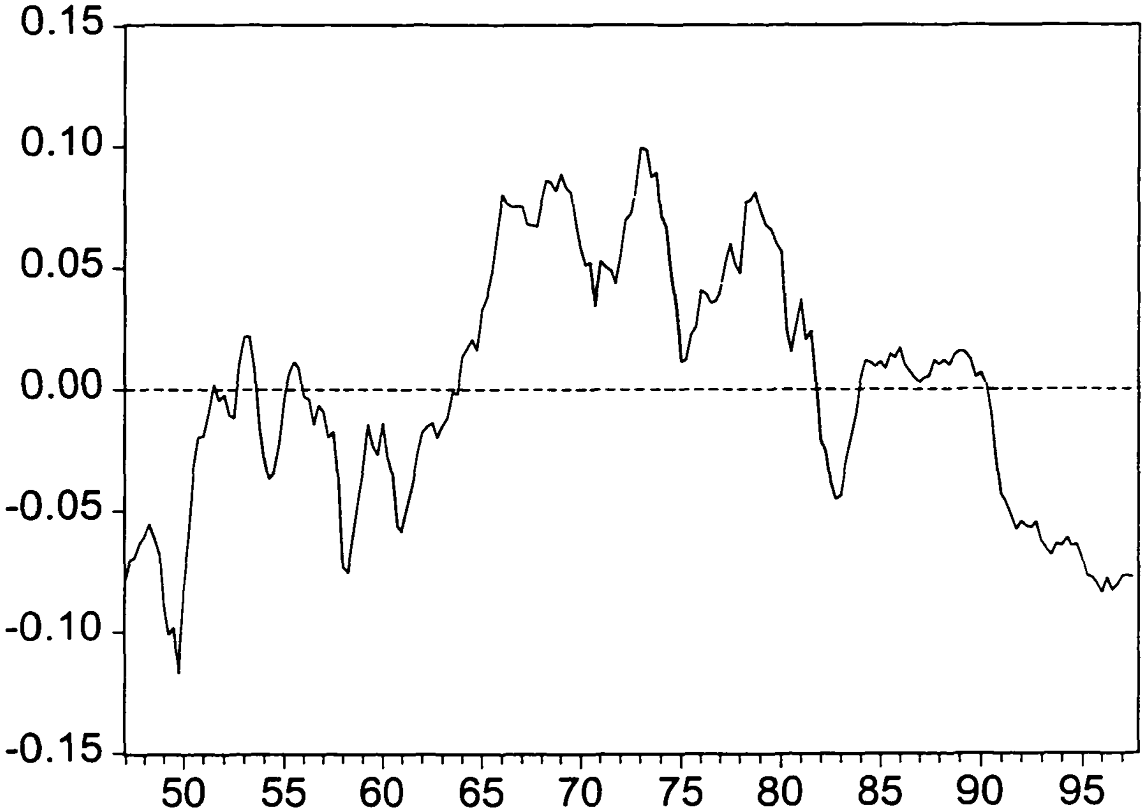

Figure 1.4 Detrended Post-war Chained Real GDP

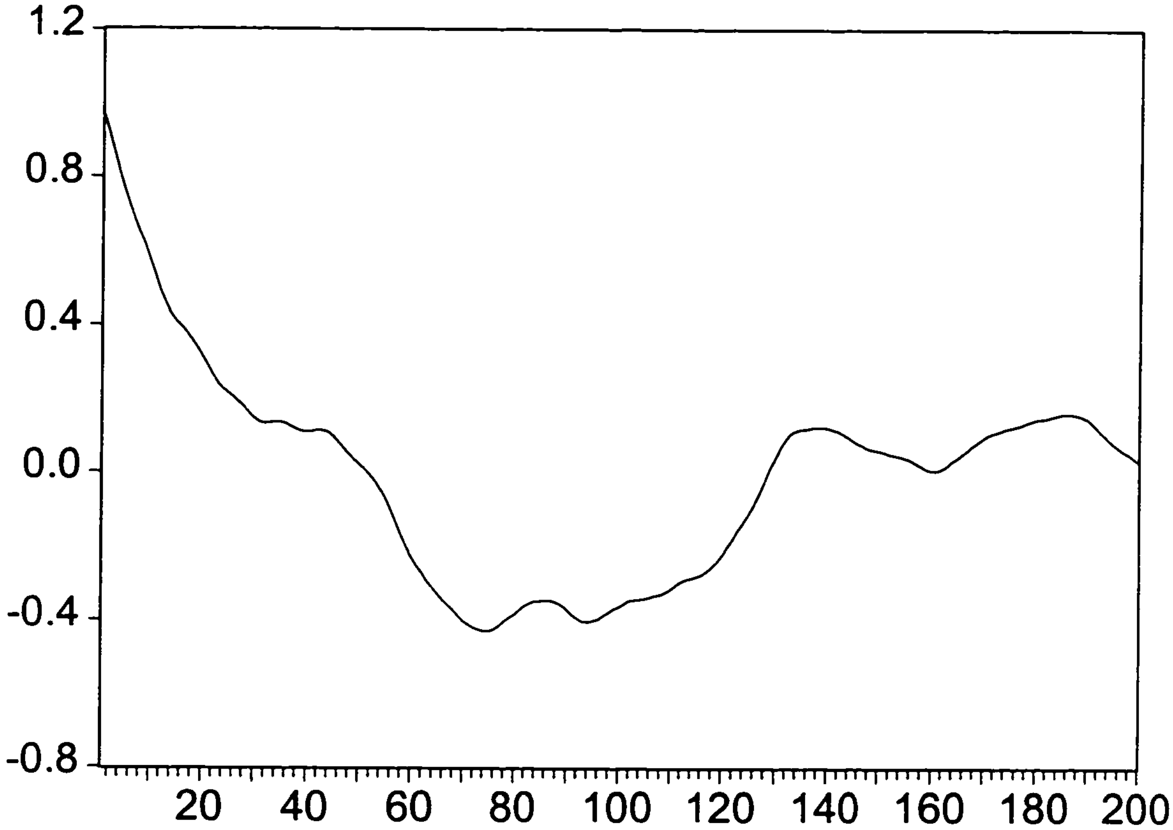

Figure 1.5 ACF of Detrended Post-war Chained Real GDP

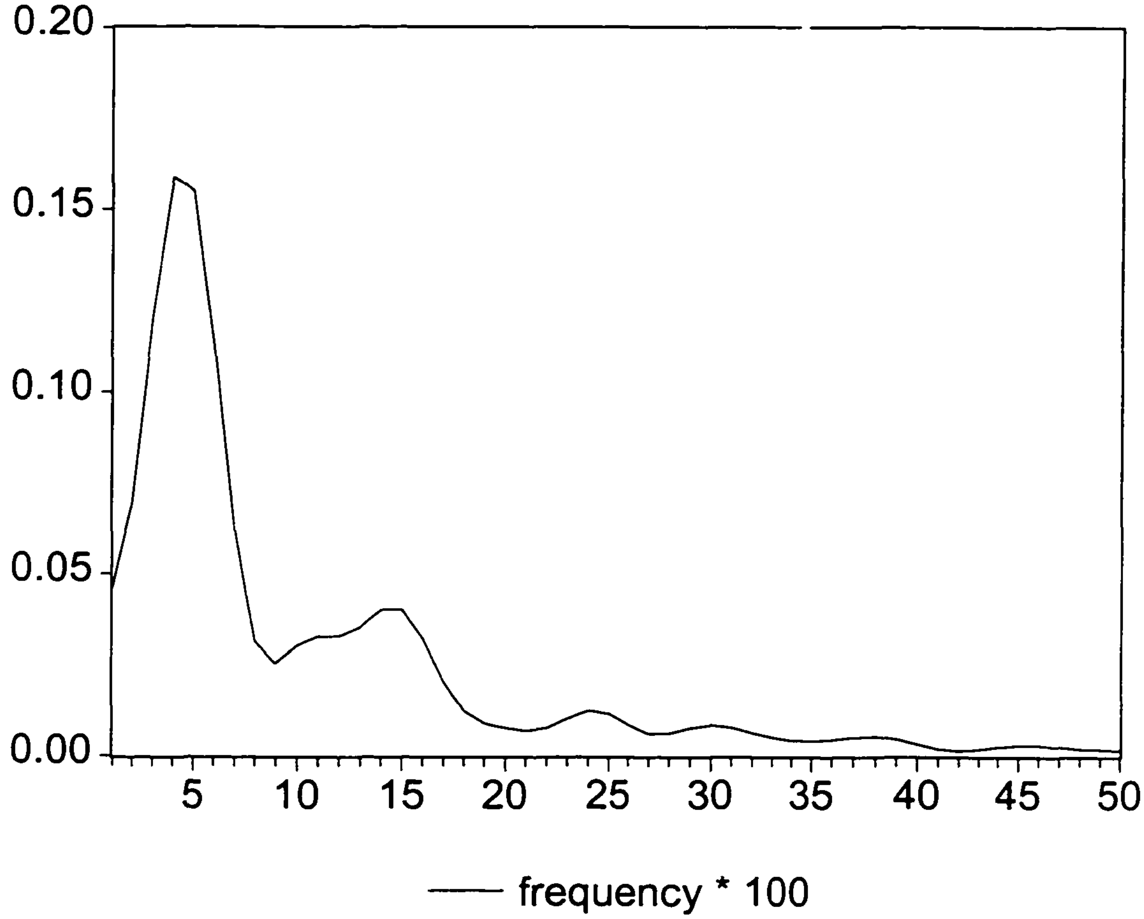

Figure 1.6 Spectrum of Detrended Post-war Chained Real GDP (Lag Window=150)

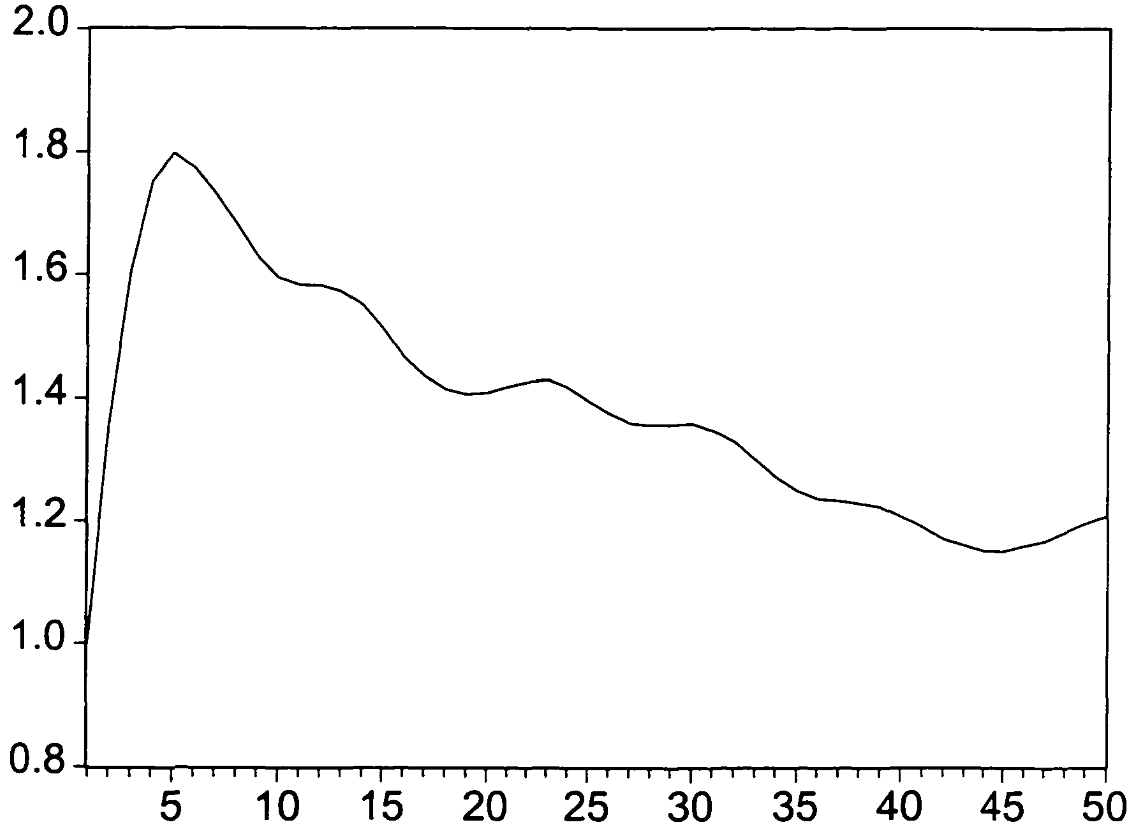

Figure 1.7 Variance Ratios for Log of Post-war Chained Real GDP Based on Cochrane's (1988) Biased Corrected Estimate

Table 1.1 Tests for a Unit Root or Trend Stationarity in U.S. Real GDP Annual Data; Maddison (1995)

Test: GS/SIC	AR Lag (4)	Test Statistic	Nominal p-value (5)	Step Dummy t-statistic
1870-1994:				
Unit root:				
Dickey-Fuller	6	-3.74	0.03	-
	1	-4.14	< .01	-
Phillips-Perron	-	-3.52	0.04	-
Perron (1)	8	-5.58	< .01	-3.74
(break in level)	1	-4.72	< .01	-2.31
Zivot & Andrews (2)	8	-6.10	< .01	-4.33
	1	-5.10	0.02	-2.82
Trend stationary:				
Leybourne	5	0.42	<.01	-
& McCabe (3)	2	0.05	0.40	-
1909-1970:				
Unit root:				
Dickey-Fuller	1 & 1	-3.43	0.06	-
Phillips-Perron	-	-2.63	0.27	-
Perron (1)	8	-4.89	< 0.01	-3.87
(break in level)	1	-4.26	0.02	-2.40
Zivot & Andrews (2)	8	-5.61	< 0.01	-4.63
	1	-4.72	0.07	-3.03
Trend stationary:				
Leybourne	4	0.47	< 0.01	-
& McCabe (3)	2	0.10	> .85	-

(1) Break in level assumed at 1929 as in Perron (1989).

(2) Break date maximizes unit root t-statistic; choose 1929 under GS & SIC.

(3) Null hypothesis is trend stationarity.

(4) GS starts with 8 lags, reducing lags until t>1.645 in absolute value,
SIC maximizes criterion of Schwarz (1978) over lags 0 to 8.

(5) Nominal p-values obtained by simulation under the null hypothesis.

(6) Exact p-values obtained by simulation with lag selection, under unit root null for unit root tests, under trend stationarity for L-M, DGP is the AR process selected by SIC for the actual data under the null.

Table 1.2. Summary Statistics for U.S. Real GDP

Growth Rates	Mean	Std Dev	Range	AR Coefficient Estimates			S E	J-B p
1870-1994	0.033	0.056	.18/-.23	0.27 *	0.00	-0.12	0.06	0.00
1870-1929	0.037	0.048	.13/-.08	-0.28 *	-0.20	-0.04	0.05	0.91
1930-1946	0.026	0.118	.18/-.23	0.90 *	-0.16	-0.49	0.08	0.93
1947-1994	0.031	0.026	.09/-.02	0.20	-0.11	-0.19	0.03	0.71
Detrended	Mean	Std Dev	Range	AR Coefficient Estimates			S E	J-B p
1870-1994	0.000	0.11	.31/-.37	1.13 *	-0.24	-0.10	0.05	0.00
1870-1929	0.029	0.082	.18/-.14	0.64 *	0.05	0.13	0.05	0.66
1930-1945	-0.100	0.222	.31/-.37	1.48 *	-0.57	-0.23	0.07	0.60
1946-1994	-0.003	0.048	.08/-.10	1.09 *	-0.24	-0.05	0.02	0.35

Notes: * denotes asymptotic t-statistic significant at .05 level.
J-B p denotes significance level of Jarque-Bera test for Normality.

Table 1.3. Monte Carlo Study of Unit Root Tests; DGP is an ARIMA(1,1,0) with Additive Outlier at 1930. Rejection Frequencies are at the Nominal 0.05 Significance Level. Lag k Selection Performed Alternately by GS and SIC.

Series Length: T=125

	Dickey-Fuller		Phillips-	Perron; 1929 Break Date				Zivot-Andrews; Search for Date				Leybourne & McCabe	
1930	GS	SIC	Perron	GS		SIC		GS		SIC		GS	SIC
Outlier	Unit Root		UR	UR	Step	UR	Step	UR	Step	UR	Step	Trend Stationary	
0	0.084	0.052	0.036	0.092	0.189	0.063	0.171	0.108	0.965	0.073	0.961	0.569	0.806
-0.1	0.076	0.045	0.054	0.147	0.256	0.115	0.239	0.115	0.959	0.066	0.961	0.549	0.794
-0.2	0.125	0.135	0.139	0.301	0.321	0.316	0.312	0.180	0.961	0.187	0.965	0.498	0.762
-0.3	0.142	0.215	0.270	0.374	0.393	0.414	0.379	0.261	0.959	0.378	0.964	0.475	0.737
-0.4	0.198	0.222	0.491	0.453	0.458	0.513	0.445	0.277	0.959	0.450	0.969	0.491	0.706

Series Length; T=62

	Dickey-Fuller		Phillips-	Perron; 1929 Break Date				Zivot-Andrews; Search for Date				Leybourne & McCabe	
1930	GS	SIC	Perron	GS		SIC		GS		SIC		GS	SIC
Outlier	Unit Root		UR	UR	Step	UR	Step	UR	Step	UR	Step	Trend Stationary	
0	0.097	0.058	0.031	0.100	0.246	0.081	0.212	0.155	0.930	0.128	0.947	0.238	0.439
-0.1	0.083	0.061	0.079	0.260	0.309	0.216	0.276	0.171	0.940	0.102	0.918	0.223	0.412
-0.2	0.228	0.286	0.299	0.448	0.425	0.508	0.402	0.364	0.923	0.366	0.931	0.182	0.346
-0.3	0.377	0.542	0.616	0.529	0.471	0.620	0.470	0.692	0.919	0.725	0.932	0.133	0.241
-0.4	0.488	0.711	0.837	0.634	0.506	0.731	0.518	0.864	0.928	0.912	0.943	0.073	0.162

Table 1.4. Monte Carlo Study of Unit Root Tests; DGP is an ARIMA(1,1,0) with an Additive Outlier Process Beginning at 1930. Rejection Frequencies are at the Nominal 0.05 Significance Level. Lag k Selection Performed Alternately by GS and SIC.

Series Length: T=125

	Dickey-Fuller		Phillips - Perron	Perron; 1929 Break Date				Zivot-Andrews; Search for Date				Leybourne-McCabe	
	GS	SIC		GS		SIC		GS		SIC		GS	SIC
Outlier	Unit Root		UR	UR	Step	UR	Step	UR	Step	UR	Step	Trend Stationary	
None	0.084	0.052	0.036	0.092	0.189	0.063	0.171	0.108	0.965	0.073	0.961	0.569	0.806
(-.2) @ 1930	0.125	0.135	0.139	0.301	0.321	0.316	0.312	0.180	0.961	0.187	0.965	0.549	0.762
(-.2) '30-'39	0.112	0.042	0.042	0.169	0.160	0.056	0.134	0.233	0.974	0.128	0.966	0.470	0.705
(-.2) 1930 on	0.050	0.033	0.033	0.056	0.272	0.040	0.256	0.227	0.975	0.187	0.974	0.577	0.813
AR(2) '30-'45	0.416	0.764	0.599	0.498	0.275	0.772	0.276	0.628	0.955	0.817	0.970	0.298	0.286
Fixed pattern	0.499	0.673	0.083	0.506	0.256	0.642	0.143	0.643	0.979	0.651	0.988	0.170	0.278

Series Length; T=62

	Dickey-Fuller		Phillips - Perron	Perron; 1929 Break Date				Zivot-Andrews; Search for Date				Leybourne-McCabe	
	GS	SIC		GS		SIC		GS		SIC		GS	SIC
Outlier	Unit Root		UR	UR	Step	UR	Step	UR	Step	UR	Step	Trend Stationary	
None	0.097	0.058	0.031	0.100	0.246	0.081	0.212	0.155	0.930	0.128	0.947	0.238	0.439
(-.2) @ 1930	0.228	0.286	0.299	0.448	0.425	0.508	0.402	0.364	0.923	0.366	0.931	0.182	0.346
(-.2) '30-'39	0.048	0.007	0.016	0.157	0.278	0.018	0.212	0.503	0.978	0.235	0.989	0.172	0.374
(-.2) 1930 on	0.060	0.018	0.033	0.084	0.412	0.053	0.347	0.483	0.953	0.411	0.963	0.319	0.501
AR(2) '30-'45	0.641	0.740	0.163	0.730	0.317	0.823	0.278	0.798	0.898	0.789	0.920	0.042	0.045
Fixed pattern	0.185	0.233	0.000	0.553	0.518	0.332	0.216	0.855	0.999	0.597	0.996	0.005	0.068

Table 1.5. Tests for a Unit Root or Trend Stationarity in Post-war Output. Quarterly Chained Real GDP; 1947.1 – 1997.3

Test: GS/SIC	AR Lag (4)	Test Statistic	Nominal p-value (5)	Exact p-value (6)
Unit root				
Dickey-Fuller	12	-1.52	0.82	0.816
	1	-2.05	0.58	0.552
Phillips-Perron	n/a	-1.81	0.69	0.560
Elliott, Rothenberg,	12	-0.84	0.90	0.621
& Stock	1	-1.50	0.59	0.660
Perron (1)	12	-3.3	0.19	0.269
(break in slope)	1	-3.96	0.05	0.056
Zivot-Andrews (2)	12	-3.31	0.54	0.628
	1	-3.99	0.14	0.173
Trend stationarity				
Leybourne	3	2.25	0.00	0.103
& McCabe (3)	1	3.17	0.00	0.021

Notes:

(1) Break in slope assumed to occur at 1973:1 as in Perron (1989).

(2) Break date maximizes unit root t-statistic; choose 1972:2 under GS & SIC.

(3) Null hypothesis is trend stationarity.

(4) GS starts with 12 lags, reducing lag until $t>1.645$ in absolute value, SIC maximizes criterion of Schwarz (1978) over lags 0 to 12.

(5) Nominal p-values obtained by simulation under the null hypothesis, as in the original articles in which these tests are described.

(6) Exact p-values obtained by simulation with lag selection, under unit root null for unit root tests, under trend stationarity for L-M, DGP is the AR process selected by SIC for the actual data under the null.

CHAPTER 2: THE NATURE OF BUSINESS CYCLE ASYMMETRY: EVIDENCE FROM A DYNAMIC FACTOR MODEL WITH REGIME SWITCHING PERMANENT AND TRANSITORY COMPONENTS

2.1 Introduction

Do economic time series permanently decrease during recessions? Recent research (Wynne and Balke (1992), Beaudry and Koop (1993), and Sichel(1994)) provides evidence which suggests that the answer is no. This implies that following a recession, output will experience above average growth. Stated another way, the "recovery" phase is indeed a recovery. In contrast, if output were to permanently decrease during a recession, there would be no recovery. Output would begin to grow from its new, lower level. In this paper, we add to the existing literature through an analysis of monthly time series which incorporates the ideas of comovement across macroeconomic time series and business cycle asymmetry.

The importance of the comovement of economic time series and business cycle asymmetry was recognized by early scholars of the business cycle. In their landmark study, Burns and Mitchell (1946) highlighted comovement as one of the two empirical regularities of the business cycle:

> ...a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle.

The other regularity of the business cycle, asymmetry, is the idea that expansions are fundamentally different than recessions. This goes back at least as far as Mitchell

(1927):

> the most violent declines exceed the most considerable advances. The abrupt declines usually occur in crises; the greatest gains occur in periods of revival,...Business contraction seems to be a briefer and more violent process than business expansion.

An often cited quote from Keynes (1936) also conveys the same message:

> ...the substitution of a downward for an upward tendency often takes place suddenly and violently, whereas there is, as a rule, no such sharp turning point when an upward is substituted for a downward tendency.

Recently, researchers have used the tools of modern time series analysis to explicitly model comovement and asymmetry. Stock and Watson (1989, 1991, 1993) estimate a linear dynamic factor model which captures the comovement across economic time series through an unobserved permanent component common to each series. The Kalman filter is used to extract the common factor which is then interpreted as a composite index of economic activity. The implied index corresponds closely to the composite index of coincident indicators developed by the Department of Commerce and indeed gives it statistical justification. Hamilton (1989) incorporates business cycle asymmetry in a univariate nonlinear model which allows the growth rate of output to be dependent on the "state" of the economy. The results from his regime switching model suggest that the economy is characterized by two states: positive growth (expansion) and negative growth (recession). Furthermore, since the asymmetry exists in the permanent component, recessions have permanent effects on the level of output.

Traditionally, comovement and asymmetry have been analyzed in isolation. In a recent paper, Diebold and Rudebusch (1996) provide empirical and theoretical support for comovement and asymmetry as important features of the business cycle and suggest that they should be analyzed simultaneously. M.-J. Kim and Yoo (1995), Chauvet (1997), and Kim and Nelson (1998a) unite the dynamic factor model and regime switching by allowing the common permanent component of Stock and Watson to undergo the type of regime switching advocated by Hamilton. M.-J. Kim and Yoo and Chauvet estimate the model via Kim's (1994) approximate maximum likelihood method, whereas Kim and Nelson employ the Gibbs sampler. They estimate new coincident indexes of economic activity which explicitly incorporate comovement and business cycle asymmetry.

In these attempts to model comovement and/or asymmetry, recessions are assumed to have permanent effects. For example, if the growth rate of a time series is subject to changes in regime, then a period of negative growth (recession) will permanently lower the level of a time series. However, recent research raises the possibility that recessions only temporarily lower output. The idea that the rate of output growth during an expansion is related to the magnitude of the preceding contraction is explicit in Friedman's (1964, 1993) "plucking" model of the business cycle:

> There appears to be no systematic connection between the size of an expansion and of the succeeding contraction...a large contraction in output tends to be followed on the average by a large business expansion; a mild contraction, by a mild expansion.

Kim and Nelson (1998b) specify a univariate nonlinear econometric model which captures Friedman's idea that recessions only temporarily affect output by allowing the transitory component of post-war real GDP to be asymmetric. The model performs remarkably well at identifying the NBER recessionary periods, and suggests that recessions are well characterized as temporary departures of output from its natural level.

It appears that neither asymmetry in the transitory component nor asymmetry in the permanent component may be ignored in modeling the business cycle. However, there has been no attempt to model both types of asymmetry simultaneously. To make up for the gap in the literature, we examine the nature of business cycle asymmetry within a dynamic factor model. While existing time series models of business cycle comovement allow a common permanent factor only, our model also allows for a common transitory factor. In addition, we examine the nature of business cycle asymmetry in terms of these two common factors. The models presented in this paper may be considered as extensions of M.-J. Kim and Yoo (1995), Chauvet (1997), and Kim and Nelson (1998a) in that they allow for a Markov switching common transitory component as well as a Markov switching common permanent component. The Markov switching common transitory component in our models potentially captures the "plucking" nature of recessions advocated by Friedman (1964, 1993) as in Kim and Nelson (1998b).

With the exception of the 1990-91 recession, which appears not to contain a significant transitory component, we find that postwar recessions are comprised of both permanent and temporary shocks. Thus, although a fraction of negative shocks to economic time series is temporary, their level is permanently lowered during a recession.

This paper is organized as follows. Section 2.2 provides a review of comovement and asymmetry in the business cycle literature. Section 2.3 presents a dynamic factor model with permanent and transitory regime switching. Section 2.4 generalizes the model by allowing the switching in the common permanent and transitory components to be independent. Section 2.5 summarizes and offers concluding remarks.

2.2 Comovement and Asymmetry in the Business Cycle Literature

The essence of the linear dynamic factor model proposed by Stock and Watson (1989, 1991, 1993) is that the comovement across economic time series can be captured by a single unobserved factor common to all the series. They analyze the four monthly coincident indicator series used to construct the Department of Commerce composite index of coincident indicators: the index of industrial production, personal income less transfer payments, manufacturing and trade sales, and employees on nonagricultural payrolls. If each series has a unit autoregressive root, it can be decomposed into a deterministic component (DT_i), and a permanent component (P_{it}):

$$Y_{it} = DT_i + P_{it}, i = 1,2,...,n,\ t = 1,2,...,T, \quad (2.2.1)$$

$$DT_i = a_i + D_i t, \quad (2.2.2)$$

$$P_{it} = \gamma_i C_t + \zeta_{it}, \quad (2.2.3)$$

where Y_{it} is the log of the i^{th} indicator series, C_t is the unobserved common permanent factor, and ζ_{it} is an idiosyncratic permanent component. If $Y_{it}, i = 1,2,\ldots,n$ are not cointegrated, the model may be written in first differences in the following manner:

$$\Delta Y_{it} = D_i + \gamma_i \Delta C_t + z_{it}, \tag{2.2.4}$$

where $z_{it} = \Delta\zeta_{it}$ is a stationary process and is approximated by:

$$\psi_i(L)z_{it} = e_{it}, \quad e_{it} \sim i.i.d.N(0,\sigma_i^2), \tag{2.2.5}$$

where the roots of $\psi_i(L) = 0$ lie outside the complex unit circle. The common permanent component, C_t, is assumed to follow an AR process in first differences:

$$\phi(L)\Delta C_t = \delta + v_t, \quad v_t \sim i.i.d.N(0,1), \tag{2.2.6}$$

where the roots of $\phi(L) = 0$ lie outside the complex unit circle and the innovation variance has been normalized to unity. The state space representation of this model is linear, and calculation of the exact Gaussian likelihood function via the Kalman filter is possible. Stock and Watson use the Kalman filter to extract an estimate of C_t which is then interpreted as a composite index of economic activity.

Hamilton (1989) analyzes business cycle asymmetry by allowing the growth rate of GNP to follow a nonlinear stationary process. If $y_t = \ln(GNP_t)$, Hamilton's model may be written as:

$$(1 - \phi_1 L - \cdots - \phi_4 L^4)(\Delta y_t - \mu_{S_t}) = e_t, \quad e_t \sim i.i.d.N(0,\sigma_e^2), \tag{2.2.7}$$

where

$$\mu_{S_t} = \mu_0 + \mu_1 S_t \text{ , } S_t = \{0,1\} . \tag{2.2.8}$$

S_t follows a 1st order Markov switching process with transition probabilities:

$$\Pr[S_t = 0 \mid S_{t-1} = 0] = q, \ \ \Pr[S_t = 1 \mid S_{t-1} = 1] = p. \tag{2.2.9}$$

Hamilton develops a nonlinear filter which facilitates calculation of the exact likelihood function for AR processes with a Markov switching mean. His results suggest that post-war quarterly GNP switches between one of two states: positive growth (expansion) and negative growth (recession).

Evidence of various types of business cycle asymmetry is abundant. Analyzing periods of increase and decline in the unemployment rate, Neftçi (1984), DeLong and Summers (1986), and Sichel (1989) provide evidence that recessions tend to be steeper and more short lived than recoveries. However, DeLong and Summers (1986) and Falk (1986) are unable to support this finding using GNP data. In addition, Sichel (1993) provides evidence that business cycle troughs are deeper than peaks are tall.

Business cycle asymmetry is also evident in the duration of expansions and recessions. Using constant transition probabilities, which imply duration independence, Hamilton (1989) and Lam (1990) find that the expected duration of an expansion is longer than that of a recession. Allowing the duration to depend on its current length, Diebold and Rudebusch (1990), Sichel (1991), Diebold, Rudebusch and Sichel (1993), Durland and McCurdy (1994), and Kim and Nelson (1998a) find that postwar contractions exhibit positive duration dependence, whereas postwar expansions do not.

The nature of business cycle asymmetry has important implications regarding the long run effect of recessions. If the permanent component is subject to changes in regime, as in Hamilton (1989), then a recession will permanently lower the level of output. Conversely, if regime switching exists only in the transitory component, then recessions will only temporarily lower the level of output.

The existing literature contains evidence of both permanent and transitory asymmetry. As mentioned earlier, Hamilton (1989), and Lam (1990) who provides a generalization of Hamilton's model, find that a model which allows the growth rate of GNP to be state dependent is able to identify turning points in economic activity which closely match the NBER recessionary periods. The regime switching dynamic factor models of M.-J. Kim and Yoo, Chauvet, and Kim and Nelson, which specify a state dependent growth rate, are also extremely successful at identifying the phases of the business cycle. This suggests that permanent variation is important in explaining recessions.

There is also evidence that the business cycle is asymmetric due to nonlinearity in the transitory component. Beaudry and Koop (1993) emphasize the significance of asymmetry in regard to the relative importance of permanent and transitory shocks. They argue that failure to allow for asymmetry results in understating the importance of positive shocks, and overstating the importance of negative shocks. In particular, they estimate impulse response functions which take into account the current depth of the recession, and conclude that the effect of a negative shock of plausible size is negligible after 12 quarters. Wynne and Balke (1992) and Sichel (1994) provide evidence that the economy grows faster immediately following a recession. A direct implication of this "peak-reverting" behavior is that declines in economic activity contain an important transitory component. This type of behavior is consistent with Friedman's (1969, 1993) "plucking" model and DeLong and Summer's (1988) "output-gaps" view of the business cycle.

Kim and Nelson (1998b) specify and estimate a univariate version of Friedman's model for output in which the source of asymmetry is regime switching in the transitory component. They employ the following unobserved components specification:

$$y_t = \eta_t + x_t \, , \qquad (2.2.10)$$

where y_t is the log of real GDP, η_t is the symmetric stochastic trend, and x_t is the transitory component. They specify asymmetry in the transitory component in the following manner:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \tau_{S_t} + u_t \, , \qquad (2.2.11)$$

$$\tau_{S_t} = \tau S_t \, , \qquad (2.2.12)$$

$$u_t \sim i.i.d.N(0, \sigma_{S_t}^2) \, . \qquad (2.2.13)$$

The cyclical component is asymmetric if $\tau \neq 0$, and $\tau < 0$ corresponds to the level of output being "plucked" down during recessionary times. As in Hamilton (1989) and Lam (1990), their model is successful at identifying the NBER recessionary periods, and the results suggest that recessions are well characterized as temporary departures of output from its natural level.

Implicit in the business cycle analysis of Burns and Mitchell is the idea that economic time series tend to move together *over the business cycle*. Indeed, a recession was said to have occurred if numerous series experienced a decline. To properly model comovement, we must first understand the nature of recessions; that is, do they permanently lower the level of output? As mentioned above, recent work by Wynne and Balke, Beaudry and Koop, and Sichel suggests that the answer, at least in part, is no. If economic time series display comovement over the business cycle and recessions only temporarily lower the level of output, this suggests the existence of a common *transitory* component. Thus we extend the dynamic factor model framework to include a common transitory factor.

2.3 A Dynamic Factor Model with Regime Switching Permanent and Transitory Common Factors: A Basic Model

To facilitate comparison with earlier work, we re-estimate the linear dynamic factor model of Stock and Watson (1989, 1991, 1993). The data used are the index of industrial production (IP), personal income less transfer payments (GMYXPQ) in billions of 1987 dollars, and manufacturing and trade sales (MTQ) in billions of 1987 dollars. DRI codes are in parentheses. The data are seasonally adjusted and the sample period is 1959.01 through 1997.01. These series comprise three of the four monthly indicator series classified by the Department of Commerce as coincident. We exclude the fourth series, employees on nonagricultural payrolls (LPNAG), since it appears to lag the business cycle (see Stock and Watson (1989, 1991, 1993) and Kim and Nelson (1998a)).

We performed the Dickey-Fuller (1979) unit root test and were unable to reject the null hypothesis of a unit root versus the alternative of trend stationarity at the 10% level for each of the three series. Also, we tested the null hypothesis that the series are not cointegrated versus the alternative that they are cointegrated using the pairwise residual based test of Engle and Granger (1987) and were unable to reject the null at the 10% level. Since the data appear to be individually integrated, but not cointegrated, the model is written as:

$$\Delta Y_{it} = D_i + \gamma_i \Delta C_t + z_{it}, \tag{2.3.1}$$

$$\Delta C_t = \delta + \phi_1 \Delta C_{t-1} + \phi_2 \Delta C_{t-2} + v_t,\ v_t \sim i.i.d.N(0,1), \tag{2.3.2}$$

$$z_{it} = \psi_{i1} z_{i,t-1} + \psi_{i2} z_{i,t-2} + e_{it},\ e_{it} \sim i.i.d.N(0,\sigma_i^2), \tag{2.3.3}$$

$$E(v_s e_{it}) = 0,\ \forall\, i,s,t, \tag{2.3.4}$$

where Y_{it} is the log of the i^{th} indicator. As in Stock and Watson (1989, 1991, 1993), we assume that the dynamics of the common and idiosyncratic components can be adequately described by a second order autoregression. The innovation variance for the common permanent factor has been normalized to unity. Stock and Watson (1991) note that with this model specification, the intercept terms, δ and $\widetilde{D} = [D_1\ D_2\ D_3]'$, are unidentified. Identification is achieved if the data are expressed in deviations from their means $(\Delta y_{it} \equiv \Delta Y_{it} - \Delta \overline{Y}_i)$. We can replace equations (2.3.1) and (2.3.2) with:

$$\Delta y_{it} = \gamma_i \Delta c_t + z_{it} , \tag{2.3.5}$$

$$\Delta c_t = \phi_1 \Delta c_{t-1} + \phi_2 \Delta c_{t-2} + v_t , \tag{2.3.6}$$

where $\Delta c_t \equiv \Delta C_t - \delta$. Writing the model in deviations from means concentrates δ and $\widetilde{D} = [D_1\ D_2\ D_3]'$ out of the likelihood function. Appendix B describes a procedure to retrieve δ and $\widetilde{D} = [D_1\ D_2\ D_3]'$.

Since the state space model is linear, the Kalman filter can be used to calculate the exact Gaussian likelihood function. Column 2 of Table 2.1 reports the maximum likelihood estimates of this model (Model 1).[1] Note that the factor loadings (γ_i) are individually significant at the 5% level.

We also consider a linear dynamic two factor model. In this framework, we decompose each series into a deterministic component, a permanent component, and a transitory component as follows:

$$Y_{it} = DT_i + P_{it} + T_{it} , \tag{2.3.7}$$

$$P_{it} = \gamma_i C_t + \zeta_{it} , \tag{2.3.8}$$

$$T_{it} = \lambda_i x_t + \omega_{it} , \tag{2.3.9}$$

[1] Throughout this paper the data are standardized so that they are distributed with zero mean and unit variance in log first differences.

where C_t is the common permanent factor, x_t is the common transitory factor, and ζ_{it} and ω_{it} are independent idiosyncratic permanent and transitory components respectively. We can write the model in deviations from means as follows:

$$\Delta y_{it} = \gamma_i \Delta c_t + \lambda_i \Delta x_t + z_{it} , \tag{2.3.10}$$

$$\Delta c_t = \phi_1 \Delta c_{t-1} + \phi_2 \Delta c_{t-2} + v_t ,\ v_t \sim i.i.d.N(0,1) , \tag{2.3.11}$$

$$x_t = \phi_1^{*} x_{t-1} + \phi_2^{*} x_{t-2} + u_t ,\ u_t \sim i.i.d.N(0,1), \tag{2.3.12}$$

where $z_{it} = \Delta\zeta_{it} + \Delta\omega_{it}$ is stationary and approximated by:

$$z_{it} = \psi_{i1} z_{i,t-1} + \psi_{i2} z_{i,t-2} + e_{it} , e_{it} \sim i.i.d.N(0,\sigma_i^2) , \tag{2.3.13}$$

and

$$E(v_r u_s e_{it}) = 0 , \forall\, i,r,s,t. \tag{2.3.14}$$

The innovations variances for both common components have been normalized to unity. In addition, the innovations variances of the three components are assumed to be mutually uncorrelated at all leads and lags.

Column 3 of Table 2.1 reports the parameter estimates for this model (Model 2). The permanent factor loadings are individually significant at the 5% level. The transitory factor loadings (λ_i) for IP and MTQ are also individually significant at the 5% level. However, the common transitory factor does not appear to be a significant explanatory variable for the personal income series (GMYXPQ). Ideally, we would test the joint null hypothesis that $\lambda_1 = \lambda_2 = \lambda_3 = 0$, in which case the model would reduce to a dynamic one factor model. However, under the null hypothesis ϕ_1^{*} and

ϕ_2^* are unidentified. This violates one of the standard assumptions of asymptotic theory which guarantees that the Wald, LR, and LM tests converge to a chi-squared random variable. Hansen (1996) provides a method for calculating critical values for the Wald, LR, and LM tests when parameters are unidentified under the null hypothesis. However, given the numerous potential nuisance parameters in the linear two factor model, his procedures are not feasible in this case.[2]

We now extend the dynamic two factor model by incorporating asymmetry in both the permanent and transitory common factors. As in the linear two factor model, each indicator variable is decomposed into a deterministic component, a permanent component, and a transitory component. With regime switching in both the permanent and transitory common factor components, equations (2.3.10) – (2.3.14) are re-written as follows:

$$\Delta y_{it} = \gamma_i \Delta c_t + \lambda_i \Delta x_t + z_{it}\,, \tag{2.3.10'}$$

$$\Delta c_t = \beta_{S_t} + \phi_1 \Delta c_{t-1} + \phi_2 \Delta c_{t-2} + v_t,\ v_t \sim i.i.d. N(0,1)\,, \tag{2.3.11'}$$

$$x_t = \tau_{S_t} + \phi_1^* x_{t-1} + \phi_2^* x_{t-2} + u_t,\ u_t \sim i.i.d. N(0,1), \tag{2.3.12'}$$

[2] To circumvent the problem of unidentified parameters under the null, we can specify the null hypothesis as $\lambda_1 = \lambda_2 = \lambda_3 = \phi_1^* = \phi_2^* = 0$. Under this hypothesis, there is no identification problem and the Wald, LR, and LM tests are asymptotically distributed as chi-squared random variables with 5 degrees of freedom. The LR statistic for this hypothesis is 23.0528 which rejects the null at the 1% significance level. This suggests that the common transitory component explains a statistically significant amount of the variation observed in the data. It should be noted that although this test has correct size asymptotically, it is not consistent against hypotheses of the form: $\lambda_1 = \lambda_2 = \lambda_3 = 0$, $\phi_1^* = a_1$, $\phi_2^* = a_2$ for a_1 and a_2 not both zero. In fact, against hypotheses of this type, the power is equal to the size for arbitrarily large samples.

$$z_{it} = \psi_{i1} z_{i,t-1} + \psi_{i2} z_{i,t-2} + e_{it} \,, e_{it} \sim i.i.d.N(0,\sigma_i^2) \,, \tag{2.3.13)'}$$

$$E(v_r u_s e_{it}) = 0 \,, \forall\, i, r, s, t \,. \tag{2.3.14)'}$$

As in M.-J. Kim and Yoo (1995), Chauvet (1997), and Kim and Nelson (1998a), we specify asymmetry in the common permanent component by allowing the growth rate of C_t to be regime dependent:[3] While δ is constant, β_{S_t} depends on whether the economy is in an expansion $(S_t = 0)$ or recession $(S_t = 1)$:

$$\beta_{S_t} = \beta_0 + \beta_1 S_t; \;\; S_t = \{0,1\}. \tag{2.3.15}$$

where S_t follows a 1st order Markov switching process with transition probabilities:

$$\Pr[S_t = 0 \mid S_{t-1} = 0] = q, \;\; \Pr[S_t = 1 \mid S_{t-1} = 1] = p. \tag{2.3.16}$$

Thus, $\delta/(1-\phi(1))$ is the long run growth rate of the common permanent component and β_{S_t} determines the deviation from this growth rate, depending on whether the economy is in an expansion or recession.

As in Kim and Nelson (1998b) we incorporate asymmetry in the transitory component by specifying:

$$\tau_{S_t} = \tau S_t. \tag{2.3.17}$$

The same unobserved state variable governs the switching of both common components. If $\tau < 0$, the transitory component is "plucked" down during

[3] Instead of specifying a state dependent intercept, Kim and Nelson allow for a state dependent mean, $\phi(L)(\Delta C_t - \delta - \mu_{S_t}) = v_t$. We allow the intercept to be state dependent purely for convenience.

recessionary times. The speed at which the plucks decay is determined by the parameters of $\phi^*(L)$.

If the state of the economy were observed, the state space model would be linear and Gaussian, and calculation of the exact likelihood function through use of the Kalman filter would be possible. The unobservability of the state, however, induces nonlinearity in the transition equation of the state space representation, and calculation of the exact likelihood function via the Kalman filter is computationally intractable. As noted by Harrison and Stevens (1976) and Gordon and Smith (1988), if there are M possible states, each iteration of the filter produces an M-fold increase in the number of states to consider. This imposes a considerable computational burden, and approximations are unavoidable. Kim (1994) proposes a method to approximate the likelihood function for state space models with Markov switching in both the measurement and transition equations. The algorithm is computationally efficient, and experience suggests that the degree of approximation is small; see Kim (1994). Presentation of the state space representation for the regime switching two factor model, and details concerning estimation are relegated to Appendix B.

The parameter estimates for this model are reported in Column 5 of Table 2.1 (Model 4). The null hypotheses that $\lambda_i = 0$ for i = 1,2, and 3 are rejected at the 5% significance level. The transitory factor loading for personal income becomes significant when we allow for asymmetric common factors. In addition, the permanent factor loadings are all individually significant. As in the linear two factor model, we would like to test the joint null hypothesis that $\lambda_1 = \lambda_2 = \lambda_3 = 0$, in which case our model would reduce to a dynamic one factor model with switching in the common permanent component. Estimates of this restricted model are reported in Column 4 of Table 1.1 (Model 3). Under the null hypothesis, ϕ_1^*, ϕ_2^*, and τ are unidentified. Although the likelihood function is increased when we allow for the common transitory factor, we cannot formally test the one factor vs. the two factor model.[4]

The expected durations of expansions and recessions are 51 and 6 months respectively.[5] It should be noted that the sum of the AR coefficients for x_t, $\phi_1^* + \phi_2^*$, is 0.6888. The common transitory component is less persistent that than of Model 2 ($\phi_1^* + \phi_2^* = 0.8149$). This corroborates the finding of Beaudry and Koop (1993) and Kim and Nelson (1998b) who demonstrate that failure to take asymmetry into account results in overstating the persistence of transitory shocks.

Figure 2.1 plots the filtered probability that a recession has occurred, whereas Figure 2.2 plots the smoothed probability, based on Kim's (1994) smoother.[6] The

---

[4] We can test the hypothesis that $\lambda_1 = \lambda_2 = \lambda_3 = \phi_1^* = \phi_2^* = \tau = 0$. The LR statistic for this hypothesis is distributed asymptotically as a chi-squared random variable with six degrees of freedom The LR statistic is 63.65, which exceeds the 1% critical value. The same caveat that applies to the hypothesis of a one factor linear model vs. a two factor linear model applies in this case as well. Asymptotically the test has power equal to size against hypotheses of the form: $\lambda_1 = \lambda_2 = \lambda_3 = 0, \phi_1^* = a_1, \phi_2^* = a_2, \tau = a_3$ for a_1, a_2, a_3 not all zero.

[5] With constant transition probabilities, the duration of expansions and recessions are calculated as $(1-p)^{-1}$ and $(1-q)^{-1}$ respectively.

[6] The filtered probability is an inference on S_t based on information up to time t: $\Pr[S_t = 1 \mid \psi_t]$, whereas the smoothed probability uses the entire sample of information: $\Pr[S_t = 1 \mid \psi_T]$

shaded areas correspond to the NBER recessionary periods. The model performs reasonably well in accounting for the NBER recessionary periods. The most notable difference between our results and those from the regime switching dynamic factor models of M.-J. Kim and Yoo (1995), Chauvet (1997), and Kim and Nelson (1998a), is that in our model neither the filtered nor the smoothed probabilities identify the 1990-91 recession.

The common transitory component, $x_{t|t}$, is plotted in Figure 2.3. Its behavior is similar to the asymmetric cyclical component of Kim and Nelson (1998b) in that it is plucked down during recessionary periods and the negative disturbances dissipate quickly. This result that output tends to grow quickly following a decline in economic activity suggests that recessions are in some part transitory. Our results thus confirm those of Wynne and Balke (1992), Beaudry and Koop (1993), and Sichel (1994) who conclude that declines in economic activity contain an important transitory component. However, there is no pluck associated with the 1990-91 recession.

Figure 2.4 plots the estimate of the common permanent component $(C_{t|t})$.[7] $C_{t|t}$ is qualitatively similar to the estimated common permanent components of M.-J. Kim and Yoo, Chauvet, and Kim and Nelson. Our results affirm their findings; namely that declines in economic activity contain an important permanent component. *This*

---

[7] $C_{t|t}$ denotes inference on C_t based on information available a time t. The procedure to extract $C_{t|t}$ based on the steady state Kalman gain is described in Appendix B. Both common factors are unitless and identified up to an arbitrary initial value.

directly implies that the level of output is permanently lowered during a recession. The parameters estimates can be used to ascertain the effect of a contraction on the level of the three indicator series. The percentage decrease of the three series in response to a six and twelve month contraction are reported in Table 2.2. The level of industrial production (IP) is permanently lowered by 2.86% and 6.67% if $S_t = 1$ for six and twelve consecutive months respectively. The answer to the question, "Do recessions permanently lower output," appears to be yes.

In spite of the model's failure to detect the most recent recession, $C_{t|t}$ drops during 1990-91. The timing of the decline is in perfect accord with the NBER's dating of the recession. In five of the six recessions in our sample, both $C_{t|t}$ and $x_{t|t}$ contract significantly. The fact that their behavior diverges during the 1990-91 recession lends credence to the popular idea that the most recent recession is qualitatively distinct from its predecessors. To identify a recession, our model requires a decrease in both common factors. A characteristic of the 1990-91 recession is that it was not followed by a period of rapid growth. For a discussion of this phenomenon, the reader is referred to Sichel (1994). Thus, even though our estimate of C_t suggests that there was indeed a period of permanent negative shocks during 1990-91, given the lack of transitory variation our model is unable to identify the most recent recession.

Our results thus indicate that declines in economic activity contain permanent and transitory components, with the exception of the most recent recession which appears to be entirely permanent.

2.4 A Dynamic Factor Model with Regime Switching Permanent and Transitory Common Factors: A Generalization

The failure of Model 4 to account for the 1990-91 recession suggests that the common permanent and transitory components do not always switch together. We thus allow them to switch independently at all points in time. In particular, we augment Model 4 to allow C_t and x_t to be driven by two independent state variables.[8] We extend equations (2.3.15)–(2.3.17) in the following manner:

$$\beta_{S_{1t}} = \beta_0 + \beta_1 S_{1t} \tag{2.4.1}$$

$$\tau_{S_{2t}} = \tau S_{2t} \tag{2.4.2}$$

Where S_{1t} and S_{2t} are independent *1st* order Markov switching processes with transition probabilities given by:

$$\Pr[S_{1t} = 0 \mid S_{1,t-1} = 0] = q_1, \;\; \Pr[S_{1t} = 1 \mid S_{1,t-1} = 1] = p_1 \tag{2.4.3}$$

and

$$\Pr[S_{2t} = 0 \mid S_{2,t-1} = 0] = q_2, \;\; \Pr[S_{2t} = 1 \mid S_{2,t-1} = 1] = p_2. \tag{2.4.4}$$

Details concerning the estimation of this model are relegated to Appendix C.

---

[8] Kim (1993) also estimates an unobserved components model for U. S. inflation with regime switching driven by two independent Markov processes.

The parameter estimates are reported in Column 6 of Table 2.1 (Model 5). Although there is little difference between most of the parameter estimates of Models 4 and 5, the behavior of S_{1t} and S_{2t} is markedly different during the 1969-70 and 1990-91 recessions.

The filtered and smoothed probabilities that C_t is in a contractionary phase are plotted in Figures 2.5 and 2.6 respectively. The common permanent component, plotted in Figure 2.9, tracks the NBER recessionary periods extraordinary well. In contrast to Model 4, the probability terms identify all of the recessionary periods in the sample, adding to the evidence that the 1990-91 recession has permanently lowered output.

The filtered and smoothed probabilities that x_t has been plucked down are plotted in Figures 2.7 and 2.8 respectively. The most notable difference in the behavior of S_{1t} and S_{2t} is that the estimated probability terms for the latter fail to identify the most recent recession. A plot of the common transitory component in Figure 2.10 conveys the same message; namely that the 1990-91 recession did not yield a high growth recovery phase. This appears to be clear evidence that the 1990-91 recession is entirely permanent.

2.5 Conclusion

While the existing literature on business cycle asymmetry focuses on asymmetry either in the permanent component (growth rate) or the transitory component separately, this paper investigates the nature of business cycle asymmetry in both components within a dynamic factor model that incorporates both a common transitory factor and a common permanent factor.

The model and empirical results in this paper provide an answer to an important question of whether recessions permanently decrease output, as recently raised by Wynne and Balke (1992), Beaudry and Koop (1993), and Sichel (1994). This paper suggests that the answer is "yes." The importance of an asymmetric common permanent component in our dynamic factor model confirms this. Our parameter estimates imply that a six month recession permanently lowers the level of industrial production by 2.86%.

However, the importance of an asymmetric common transitory component also suggests that following a recession, output will experience above average growth: *i.e.* a fraction of the negative shocks to output associated with the recession will decay. This asymmetric transitory component potentially captures the plucking nature of recessions advocated by Friedman (1964, 1993) as in Kim and Nelson (1998b). Recently, Kim and Nelson (1998a), based on a dynamic factor model with only a permanent common component, provide evidence of positive duration dependence for recessions but not for booms. The significance of the asymmetric transitory component or the plucking term in our model potentially explains the findings of Kim and Nelson (1998a).

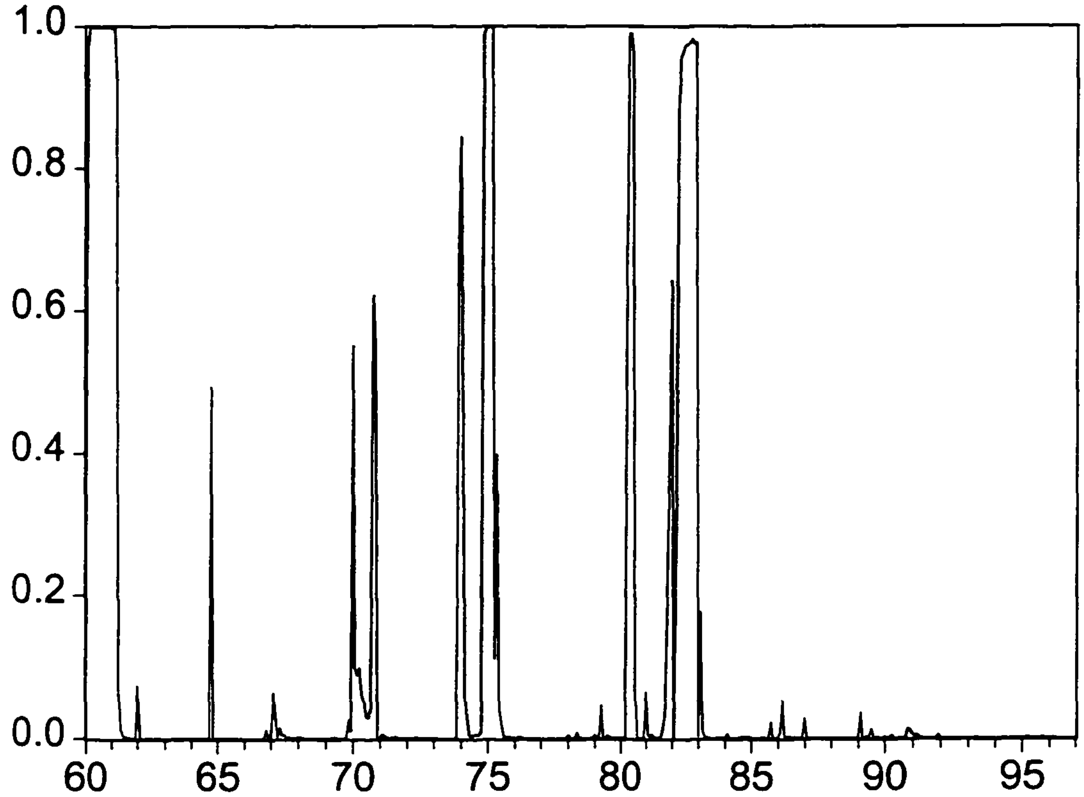

Figure 2.1 Filtered Probability of a Recession, Model 4

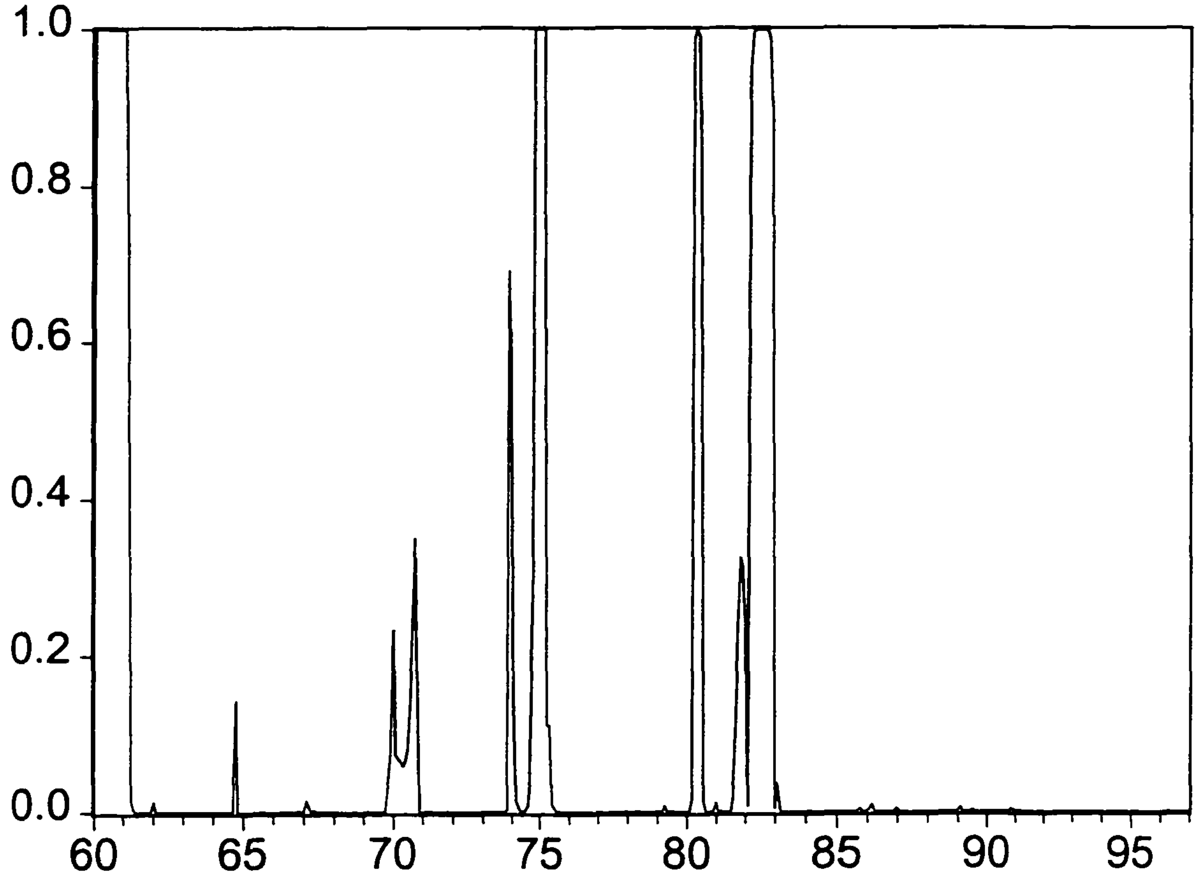

Figure 2.2 Smoothed Probability of a Recession, Model 4

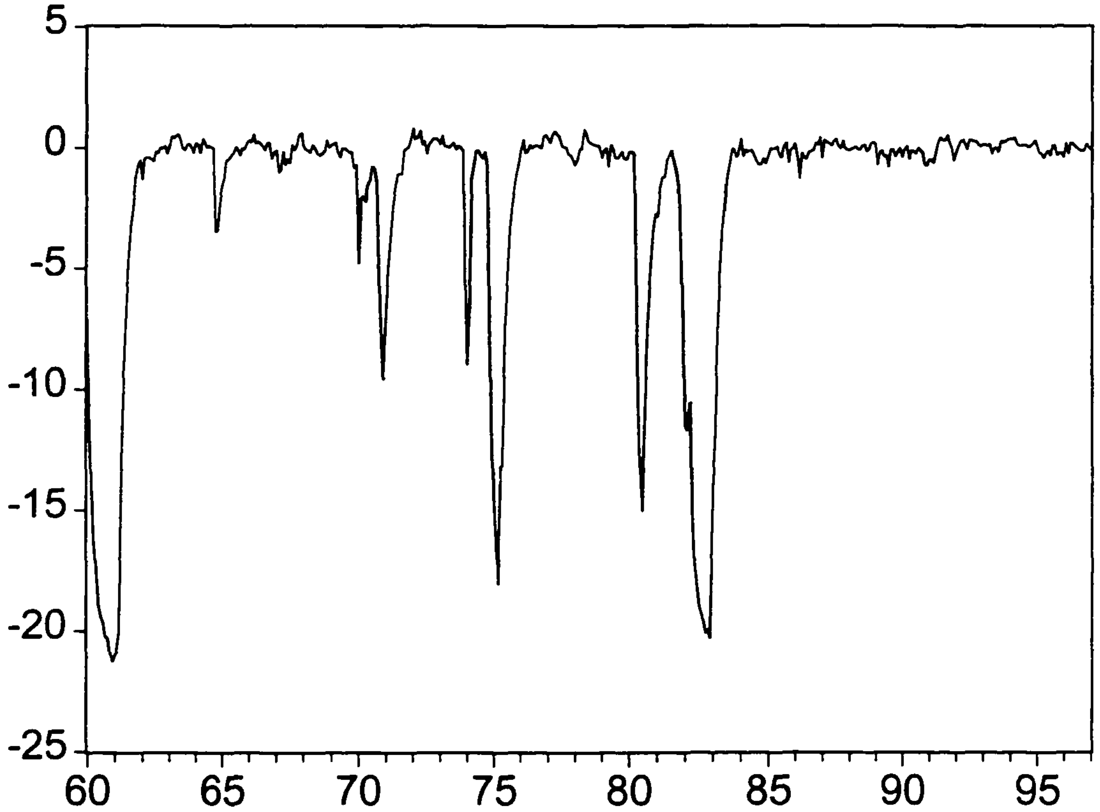

Figure 2.3 Common Transitory Component, Model 4

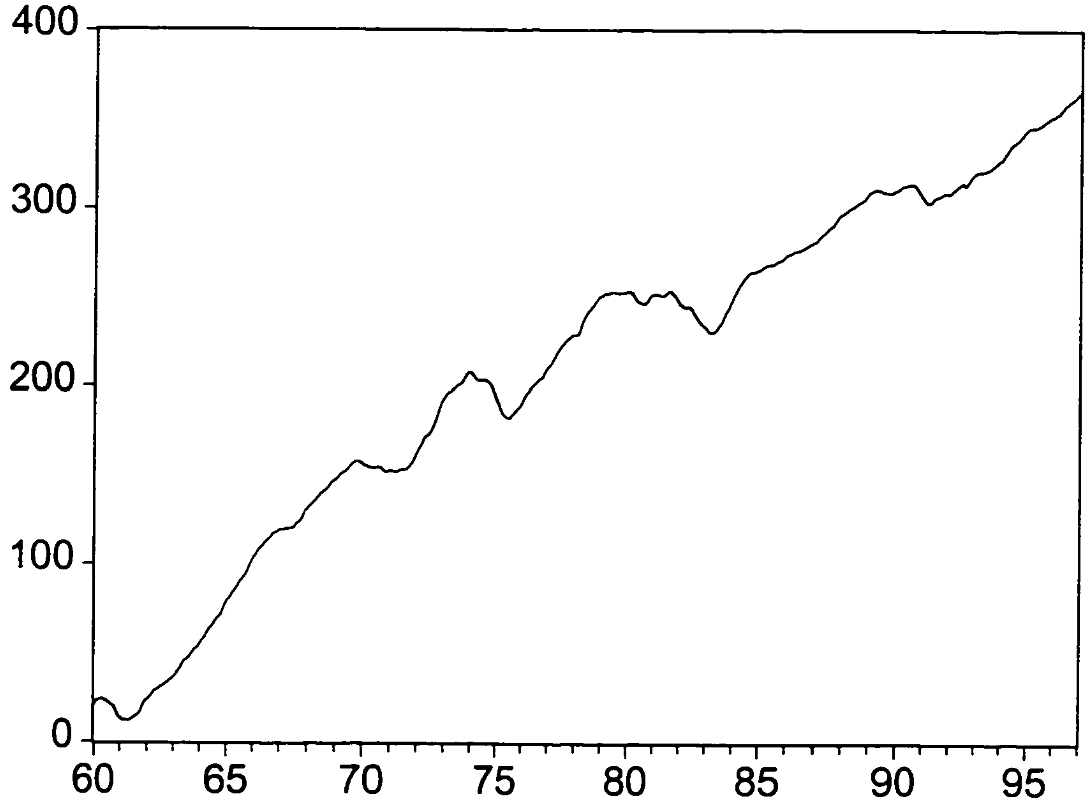

Figure 2.4 Common Permanent Component, Model 4

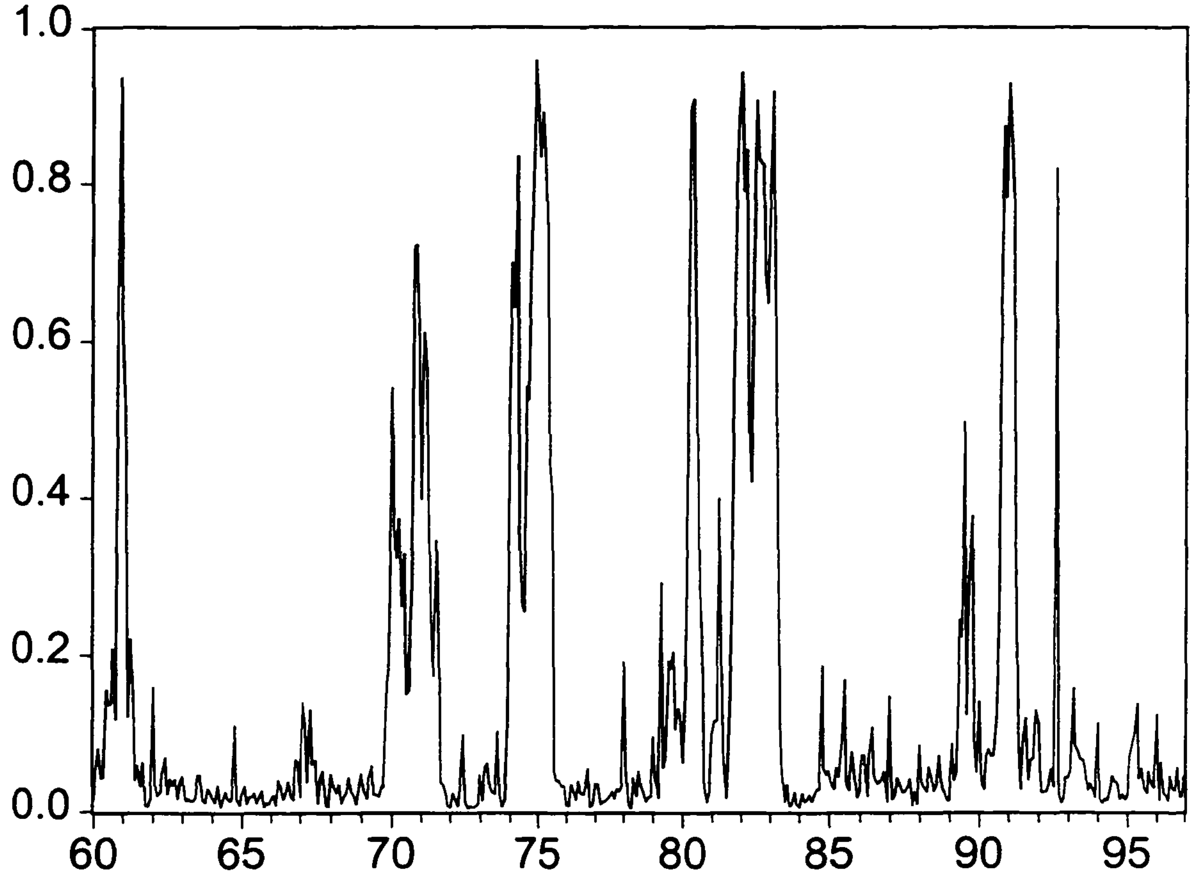

Figure 2.5 Filtered Probability that C_t is contracting, Model 5

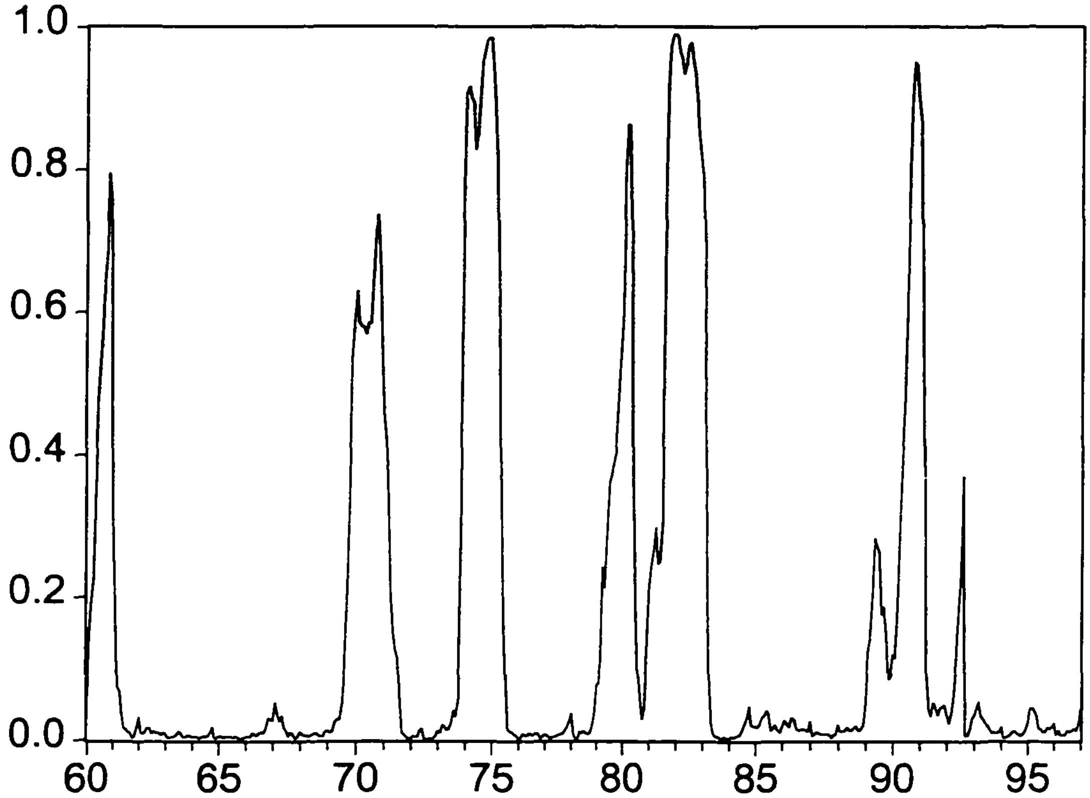

Figure 2.6 Smoothed Probability that C_t is contracting, Model 5

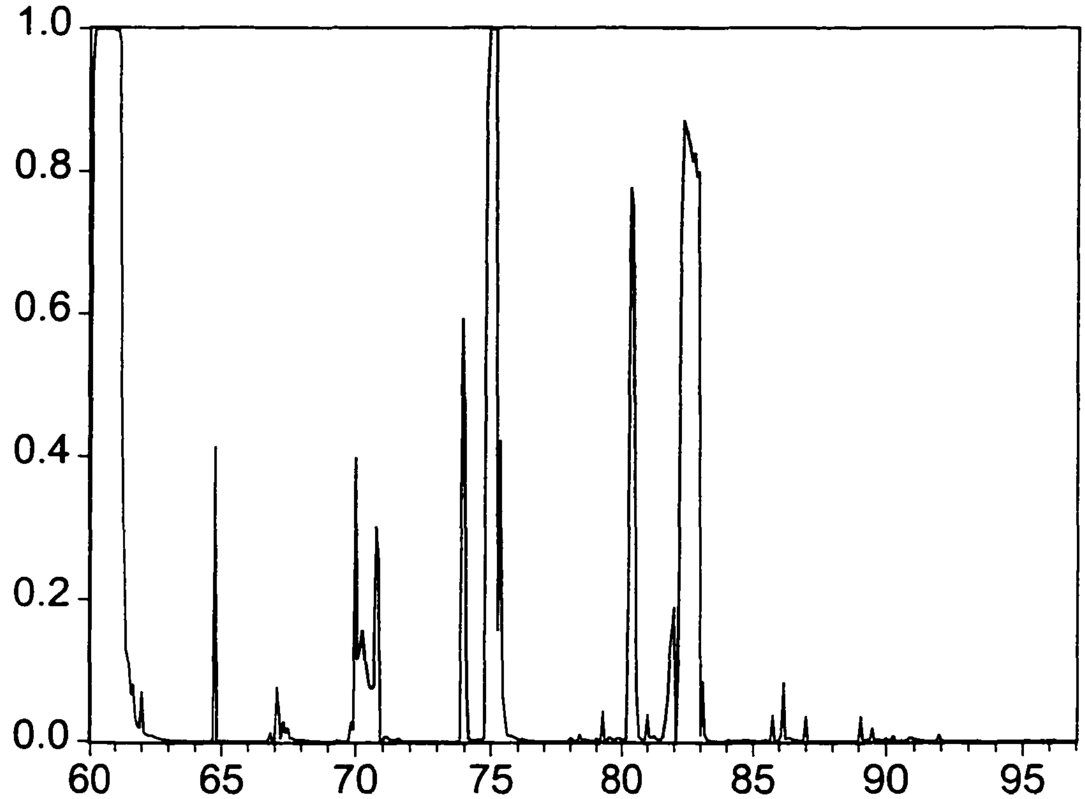

Figure 2.7 Filtered Probability that x_t is contracting, Model 5

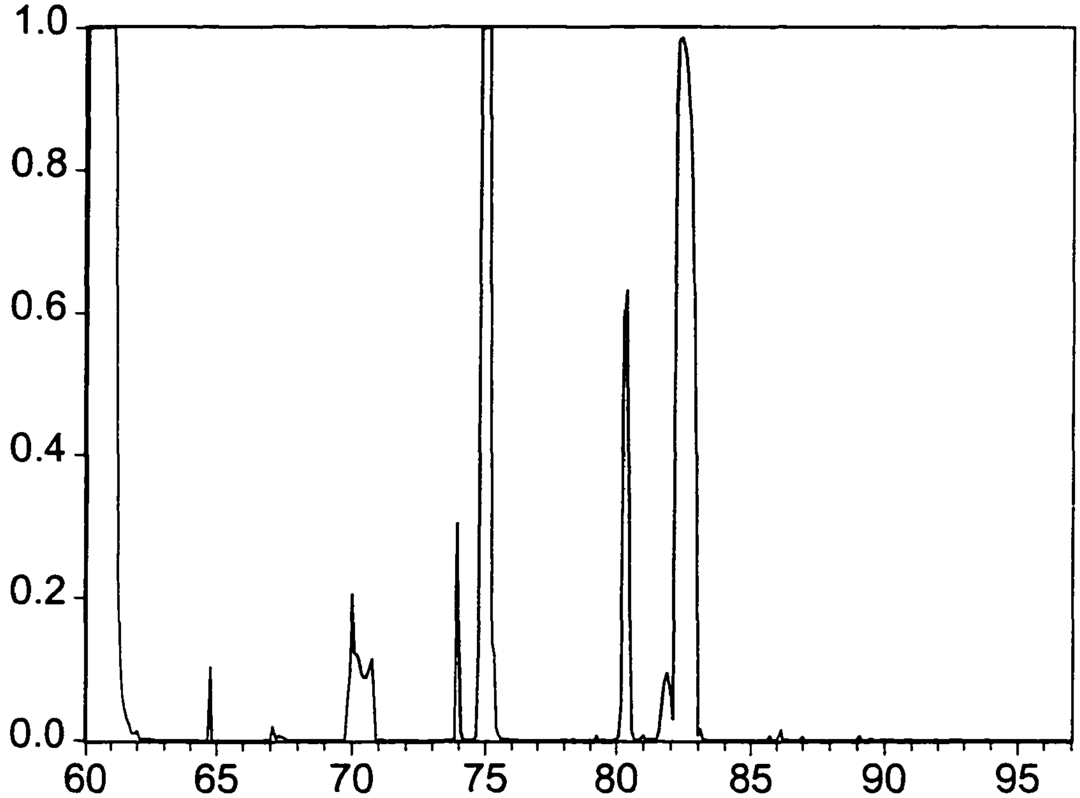

Figure 2.8 Smoothed Probability that x_t is contracting, Model 5

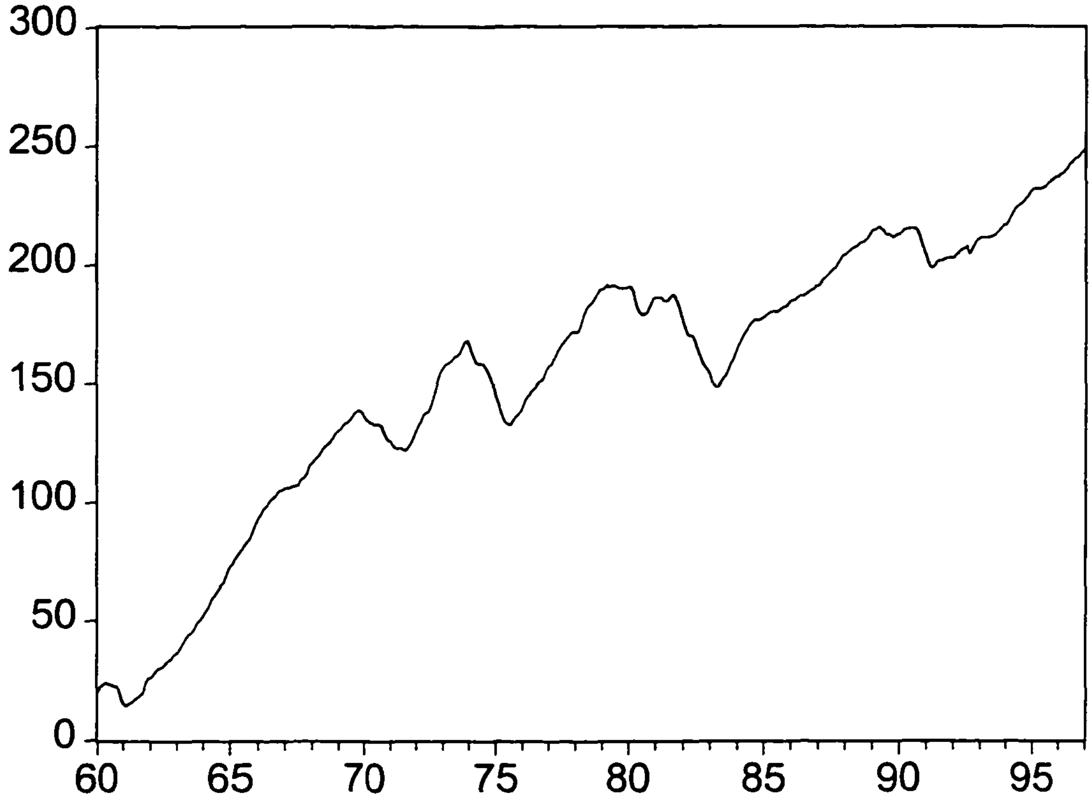

Figure 2.9. Common Permanent Component, Model 5

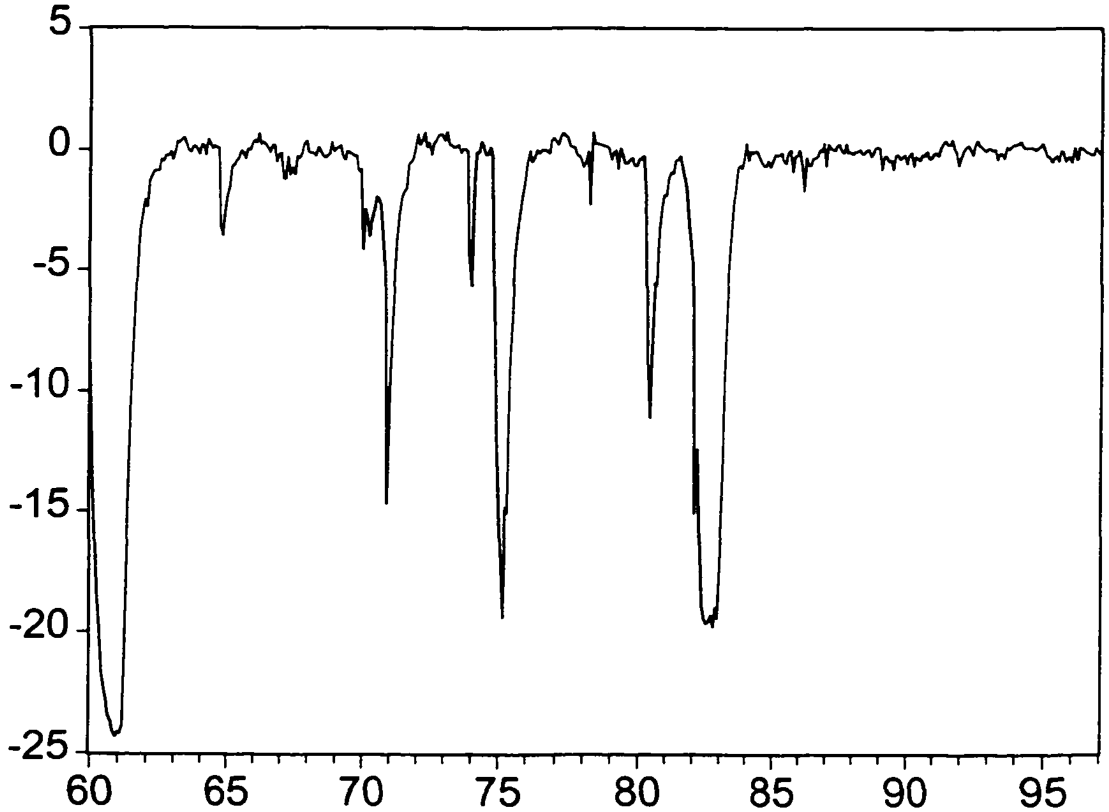

Figure 2.10 Common Transitory Component, Model 5

Table 2.1. Maximum Likelihood Estimates

Parameters	Model 1	Model 2	Model 3	Model 4	Model 5
q	----	----	0.9869	0.9801	----
			(0.0082)	(0.0070)	
p	----	----	0.4267	0.8320	----
			(0.1710)	(0.0555)	
q_1	----	----	----	----	0.9812
					(0.0106)
p_1	----	----	----	----	0.8751
					(0.0660)
q_2	----	----	----	----	0.9829
					(0.0067)
p_2	----	----	----	----	0.8390
					(0.0586)
ϕ_1	0.5173	0.5175	0.4209	0.6875	0.6769
	(0.0791)[a]	(0.0898)	(0.0876)	(0.2017)	(0.2170)
ϕ_2	0.0464	0.0786	0.0985	0.0455	-0.1145
	(0.0721)	(0.0779)	(0.0756)	(0.1472)	(0.0735)
ϕ_1^*	----	1.0316	----	0.6563	0.6388
		(0.0879)		(0.0832)	(0.0915)
ϕ_2^*	----	-0.2167	----	0.0325	0.0731
		(0.0743)		(0.0674)	(0.0733)
ψ_{11}	0.0369	-0.3299	-0.0401	0.1227	0.1654
	(0.0702)	(3.2524)	(0.1685)	(0.1365)	(0.1590)
ψ_{12}	-0.0003	0.2657	-0.0004	-0.0038	-0.0068
	(0.0013)	(0.7354)	(0.0034)	(0.0084)	(0.0132)
ψ_{21}	-0.0923	-0.1788	-0.0757	-0.1288	-0.1518
	(0.0547)	(0.0711)	(0.0566)	(0.0570)	(0.0576)
ψ_{22}	0.0160	-0.0056	0.0210	-0.0041	-0.0058
	(0.0489)	(0.0218)	(0.0525)	(0.0037)	(0.0044)
ψ_{31}	-0.4255	-0.3945	-0.4017	-0.3971	-0.4057
	(0.0561)	(0.0523)	(0.0563)	(0.0533)	(0.0539)
ψ_{32}	-0.0453	-0.0389	-0.0403	-0.0394	-0.0412
	(0.0119)	(0.0103)	(0.0113)	(0.0106)	(0.0109)
γ_1	0.6946	0.6133	0.6457	0.3316	0.3150
	(0.0484)	(0.0719)	(0.0502)	(0.0876)	(0.0813)
γ_2	0.5073	0.5667	0.4446	0.3344	0.3240
	(0.0418)	(0.0553)	(0.0421)	(0.0863)	(0.0834)
γ_3	0.4890	0.4403	0.4317	0.2309	0.2254
	(0.0372)	(0.0483)	(0.0381)	(0.0648)	(0.0646)
λ_1	----	0.6134	----	0.4076	0.4030
		(0.0676)		(0.0742)	(0.0721)

λ_2	----	-0.0629	----	0.1551	0.1542
		(0.0827)		(0.0510)	(0.0445)
λ_3	----	0.1560	----	0.2487	0.2477
		(0.0676)		(0.0486)	(0.0475)
σ_1	0.5593	0.0011	0.5195	0.4299	0.4489
	(0.0426)	(0.0403)	(0.0540)	(0.0554)	(0.0600)
σ_2	0.7953	0.7134	0.8078	0.7950	0.7814
	(0.0321)	(0.0484)	(0.0337)	(0.0340)	(0.0324)
σ_3	0.7318	0.7610	0.7490	0.7534	0.7476
	(0.0300)	(0.0287)	(0.0317)	(0.0285)	(0.0288)
β_0	----	----	0.0732	0.0628	0.2171
			(0.0553)	(0.0556)	(0.0877)
β_1	----	----	-3.350412	-0.6242	-1.7270
			(0.7589)	(0.2251)	(0.3920)
τ	----	----	----	-6.4700	-6.9620
				(1.1430)	(1.3673)
$\ln L$	-472.2802	-460.7538	-881.0622	-849.2369	-850.1289

[a] Standard errors of the parameters estimates are reported in parentheses.

Table 2.2 Permanent Decreases in the Indicators Due to a Recession

Series	Length of Recession	
	6 months	12 months
Industrial Production	2.86%	6.67%
Personal Income Less Transfer Payments	1.60%	3.63%
Manufacturing and Trade Sales	2.37%	5.54%

CHAPTER 3: INFERENCE ON UNIT ROOTS AND TREND BREAKS IN MACROECONOMIC TIME SERIES

3.1 Introduction

The most damaging criticism of the hypothesis advanced by Nelson and Plosser (1982), that U.S. output contains a unit root, has come through the allowance of structural change under the alternative hypothesis of trend stationarity. This was originally due to Perron (1989) and Rappoport and Reichlin (1989) who argued that Nelson and Plosser had overstated the frequency and magnitude of permanent shocks by failing to allow for a one time structural change under the alternative hypothesis. Perron showed that the real GNP series used by Nelson and Plosser is no longer consistent with the unit root hypothesis if a change in level, occurring at 1929, is considered. Perron's conclusion is that from 1909 to 1970, there is only one permanent shock, a negative one, and the rest of the variation in output is transitory around a time trend.

In Perron (1989), the date of the trend break, 1929, was assumed to be known *a priori*. This drew criticism originally from Christiano (1992) who suggested that Perron's results may be tainted by the assumption that the break date was known. Using a bootstrap procedure, he demonstrated that if the break date is allowed to be data dependent, then the critical values are much larger (in absolute value) than those tabulated by Perron. Zivot and Andrews (1992) and Banerjee et. al (1992) derived the limiting distribution of the unit root statistic when the break date is endogenized.

Zivot and Andrews (1992) demonstrate that Perron's conclusion that U.S. GDP is stationary around a broken time trend still holds once critical values are adjusted to reflect estimation of the break date.

Since Perron, the literature has been flooded by papers which study the asymptotic distribution of unit root and/or trend break statistics under various methods for selecting the break date. This paper adds to the literature by deriving the asymptotic distribution of statistics on structural change coefficients, as well as statistics testing the joint null hypothesis of a unit root and no structural change. The latter potentially offer an increase in power over statistics which just test the unit root null. We then apply our results to the Maddison (1995) annual U. S. real GDP series, and post-war quarterly chained real GDP.

This paper is organized as follows. Section 3.2 reviews the literature on testing for unit roots and trend breaks. Section 3.3 presents and derives the asymptotic distribution of our test statistics. Section 3.4 analyzes finite sample size and power. Section 3.5 applies our results to U.S. GDP. Section 3.6 summarizes and offers concluding remarks.

3.2 Testing for Unit Roots and Trend Breaks: A Brief Review of the Literature

Scattered throughout the literature is a plethora of results on the asymptotic distribution of unit root and structural change statistics when the break date is endogenized. In this section, we review these results for models which allow for (at most) one break in trend, and point out what has yet to be done. We divide the cases into trending and non-trending data.

3.2.1 Non-trending data

For non-trending data, the null hypothesis is a driftless unit root process with or without break, and the alternative is a stationary process with a one time change in mean. There are two methods of modeling trend breaks in the literature. The additive outlier (AO) approach models the break as an abrupt change, while the innovational outlier (IO) approach allows the break to occur gradually. Since most of the empirical work has used the (IO) approach, we concentrate on this method. For a detailed discussion of modeling innovational and additive outliers, the reader is referred to Vogelsang and Perron (1994). In general, all statistics for non-trending data are asymptotically invariant to a mean shift under the null hypothesis. We thus present the following null hypothesis without a break:

$$H_0 : y_t = y_{t-1} + u_t \tag{3.1}$$

where $u_t = \psi^*(L)e_t$; $\psi^*(L) = (1-\rho L)\psi(L)$; $\psi(L) = A(L)^{-1}B(L)$; where $e_t \sim iid(0,\sigma^2)$ and $A(L)$ and $B(L)$ are p^{th} and q^{th} order lag polynomials with roots strictly outside the unit circle. The alternative hypothesis allows for a one time change in mean and is as follows:

$$H_1 : \ y_t = c + \psi(L)(\theta DU(T_B)_t + e_t), \tag{3.2}$$

where $DU(T_B)_t = 1$ if $t > T_B$ and 0 otherwise; $DU(T_B)_t$ is the "step dummy" capturing a level shift at time T_B (the break date) and θ represents the immediate change in mean under the alternative hypothesis. For this model, the test regression is:

$$y_t = \hat{c} + \hat{\theta}DU(\hat{T}_B)_t + \hat{\rho}y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t , \tag{3.3}$$

where $\hat{T}_B$ is the estimated break date and $\hat{k}$ is the estimated lag length.

Following the theoretical treatment of Said and Dickey (1984), k lagged differences are included in the regression equation to account for serial correlation in the innovation sequence. Said and Dickey prove that if k diverges as T diverges, but at a slower rate, then the asymptotic distribution of the ADF test is unaffected. For unit root tests which allow for structural change at an unknown change point, there does not exist a proof that such a result is valid. In subsequent theoretical derivations, we shall assume that the errors are i.i.d. which simplifies the presentation of the results. We follow Zivot and Andrews (1992) and conjecture that adding k lags to the regression will correct for serial correlation.

As another matter, the correct number of lagged terms to include in the regression equation is unknown and must be chosen by the researcher. Choosing k too small results in a size bias, while choosing k too large results in a loss of power. In practice, certain data dependent methods for selecting k lead to an increase in power over fixing k as in Said and Dickey (1984) (unless of course you happen to choose the correct value of k). For standard ADF regressions, Hall (1994) proves that a number of such data based procedures leave the asymptotic distribution of the unit root statistic unaffected when the error terms follow a pure AR(p) process. Ng and Perron (1995) extend Hall's results to the ARMA(p,q) case. Among the methods analyzed are a general to specific (GS) strategy and the Schwartz information criterion (SIC).

As long as the maximum lag in the selection set is allowed to grow appropriately with the sample size, both methods are shown to have zero probability of underfitting as the sample size diverges. This implies that the asymptotic critical values, which assume that k is known, are valid under such data dependent methods for selecting k. While such a result is likely to hold for unit root tests with structural change at an unknown point, a proof is likely to be quite involved. Again we conjecture that such a result exists, and in the subsequent empirical application, we shall employ both GS and SIC.

It should also be noted that for a particular regression, the lag length and break date are determined simultaneously. This will influence the finite sample performance of the test statistics. The appropriate method used to choose T_B is context specific. If rejection of the unit root hypothesis is desired, then $\inf t_\rho$ is the appropriate statistic. However, if one is just concerned with the dating of structural change, then choosing the break date to maximize some function of $\hat{\theta}$ is appropriate.

Table 3.1 presents the relevant statistics from regression (3.3), and their origin. Blank spaces indicate what has not yet be done. Perron and Vogelsang (1992) derive the asymptotic distribution of the unit root statistic where the break date is chosen to minimize the unit root statistic. This is denoted as $\inf t_\rho$. They demonstrate that the additive outlier approach is asymptotically equivalent to the innovational outlier approach. They also consider the distribution of the unit root statistic when the break is chosen to minimize the one sided t-test of no structural change. This statistic is denoted as $t_{\rho,\inf(\theta)}$. In general, when a dummy variable statistic is used to choose the break date, the asymptotic equivalence of unit root statistics between the AO and IO approaches does not hold.

Incorporating *a priori* knowledge of the sign of the break date can lead to an increase in power. Perron and Vogelsang demonstrate that $t_{\rho,\inf(\theta)}$ has greater power than $\inf t_\rho$ when the break date is negative. A similar result holds for $t_{\rho,\sup(\theta)}$ when $\theta > 0$.

The literature also contains some distributional results for tests statistics concerning structural change coefficients. Perron and Vogelsang (1992) derive the asymptotic distribution of the mean shift statistic, but critical values are not reported. Vogelsang (1997), modeling the break as an additive outlier, derives the asymptotic distribution of the mean-Wald, exp-Wald, and sup-Wald tests of the hypothesis of no structural change for *I(0)* and *I(1)* data. The mean-Wald and exp-Wald tests cannot be used to estimate the break date, whereas the sup-Wald test can. These are extensions of the optimal tests considered by Andrews (1993) and Andrews and Ploberger (1994) for deterministically and stochastically trending data. However, the optimality properties do not carry over to trending data. Vogelsang just considers the 2-sided Wald test that $\theta = 0$. Also of interest are the 1-sided t-tests that $\theta = 0$, which may lead to higher power if the sign of the break date is known.

Two statistics in this context have not yet been computed. The first concerns inference on θ when T_B is chosen to minimize t_ρ. Second is the Wald test of the joint null that $\rho = 1$ and $\theta = 0$. We denote this statistic as $\sup Wald_{\theta,\rho}$. This may offer an increase in power over the $\inf t_\rho$ and $\sup Wald_\theta$ statistics which do not explicitly test a subset of the null hypothesis.

3.2.2 Trending data

For trending data, three different alternative hypotheses have been considered. The first, labeled Model A by Perron (1989) allows for a change in level under the alternative hypothesis. Model B allows for a change in the growth rate under the alternative, and Model C allows for both types of structural change. In general, all statistics for trending data are asymptotically invariant to a level shift under the null, but not to a change in slope. Thus, statistics for Model B and Model C will have different limiting distributions depending on whether a change in growth is allowed under the null. However, as pointed out by Vogelsang and Perron (1994), for changes in growth of the size typically encountered in practice, the no break asymptotics provide a better approximation to the finite sample distribution of the unit root statistics. We will thus present the models without a change in level or growth under the null.

All three models have the common null hypothesis:

$$\mathrm{H}_0 : y_t = \mu + y_{t-1} + u_t , \tag{3.4}$$

where $\{u_t\}$ obeys the restrictions in (3.1). The three alternative hypotheses can be written as follows:

$$\mathrm{H}_1^A : \ y_t = c + \mu t + \psi(L)(\theta DU(T_B)_t + e_t) , \tag{3.5}$$

$$\mathrm{H}_1^{\mathrm{B}}: \ y_t = c + \mu t + \psi(L)(\gamma DT(T_B)_t + e_t) \ , \tag{3.6}$$

and

$$\mathrm{H}_1^{\mathrm{C}}: \ y_t = c + \mu t + \psi(L)(\theta DU(T_B)_t + \gamma DT(T_B)_t + e_t) \ . \tag{3.7}$$

The "ramp" dummy $DT(T_B)_t$ is $t - T_B$ if $t > T_B$ and 0 otherwise, and γ is the immediate change in growth allowed under the latter two alternatives. The corresponding test regressions are:

$$y_t = \hat{c} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\beta} t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t \ , \tag{8}$$

$$y_t = \hat{c} + \hat{\gamma} DT(\hat{T}_B)_t + \hat{\beta} t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t \ , \tag{9}$$

and

$$y_t = \hat{c} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\gamma} DT(\hat{T}_B)_t + \hat{\beta} t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t \ . \tag{10}$$

For the sake of clarity, we shall discuss the origin of the statistics for all three models separately. Table 3.2 contains a description of the Model A statistics.

Choosing the break date to minimize the unit root statistic, Zivot and Andrews (1992), Banerjee et. al (1992), and Perron (1997) derive the distribution of unit root statistic, $\inf t_\rho^A$, under the no break null with innovational outliers. Modeling the break as an additive outlier, Vogelsang and Perron (1994) derive $\inf t_\rho^A$ with no break under the null. Vogelsang and Perron (1994) and Perron (1997) also derive the unit root statistic when T_B is chosen via a statistic on $\hat{\theta}$. We will generically refer to this as $t^{A}_{\rho,Wald(\theta)}$, even though the break date is usually chosen to maximize or minimize the 1-sided t-test that $\theta = 0$.

Banerjee et. al (1992) derive the Wald test that $\theta = 0$, denoted $\sup Wald^{A}_{\theta}$. Also of interest are the 1-sided t-tests of the same hypothesis.

As in the case of non-trending data, neither $Wald^{A}_{\theta.\inf(\rho)}$ nor $\sup Wald^{A}_{\theta.\rho}$ have yet been considered. The former is appropriate when one performs the Zivot-Andrews Model A unit root test, and then wishes to perform inference on θ. As mentioned before, the latter may offer an increase in power over either $\inf t^{A}_{\rho}$ or $\sup Wald^{A}_{\theta}$.

Table 3.3 presents analogous results for Model B. Since they basically mirror Table 3.2, we forgo a discussion.

Model C results are presented in Table 3.4. Given that there are 2 structural change coefficients, there are many more cases to consider. To conserve space, we shall primarily focus on what has not yet been done. Of interest is the distribution of the unit root statistic, when the break date is chosen to maximize the joint Wald test that $\theta = \gamma = 0$, denoted $t^{C}_{\rho.Wald(\theta.\gamma)}$. There is also the joint Wald test that $\theta = \gamma = 0$ when the break date minimizes the unit root statistic, $Wald^{C}_{\theta.\gamma.\inf(\rho)}$. Although Vogelsang derives the mean-Wald, exp-Wald, and sup-Wald tests of the hypothesis that Model C contains no level shift or a change in growth, again for both *I(0)* and *I(1)* data, the individual 1 and 2-sided tests are of interest. Finally, there is the joint Wald test that $\rho = 1$ and $\theta = \gamma = 0$.

In Section 3.3, we shall catalog the distributions of test statistics for no structural change, and derive the distributions of test statistics for the joint null hypothesis that there is a unit root without a break in trend.

3.3 Asymptotic Distribution of the Test Statistics

In this section, we derive the asymptotic distributions of structural change statistics, as well as the joint distributions of statistics concerning the largest autoregressive root and structural change coefficients. The latter potentially offer a gain in power over tests which do not explicitly test the unit root hypothesis. In the theorems to follow, we restrict the innovation sequence to be i.i.d., but the results remain valid in the presence of ARMA(p,q) errors. We consider non-trending and trending data separately.

3.3 Non-trending data

Following Zivot and Andrews (1992) and Banerjee et. al (1992) we specify a no break null hypothesis and innovational outliers. Recall the null hypothesis and test regression:

$$\mathrm{H}_0: \; y_t = y_{t-1} + u_t, \qquad (3.1)'$$

and

$$y_t = \hat{c} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t \,. \qquad (3.3)'$$

Let $\lambda \equiv \frac{T_B}{T}$ be the break fraction. For all the results which follow, we assume that λ remains constant as $T \to \infty$.

We first consider 4 different statistics to test the null hypothesis that $\theta = 0$. Let $\sup Wald_\theta$ and $\sup t_{|\theta|}$ be the 2-sided tests where λ is chosen to maximize the Wald statistic, and the absolute value of the t-statistic respectively. Also, let $\sup t_\theta$ and $\inf t_\theta$ be the 1-sided tests which maximize and minimize the t-statistic respectively. The latter should be used if one has *a priori* knowledge of the sign of θ.

Following Zivot and Andrews, we can characterize the asymptotic distributions of these statistics in terms of projection residuals. Let $DU^*(\lambda, r)$ be the projection residual from the continuous time regression:

$$DU(\lambda, r) = \hat{\alpha}_0 + \hat{\alpha}_1 W(r) + DU^*(\lambda, r);$$

where $DU^*(\lambda, r) = 1$ if $r > \lambda$ and 0 otherwise, and $W(r)$ is standard Browning motion. That is, $\hat{\alpha}_0$ and $\hat{\alpha}_1$ solve

$$\min_{\alpha} \int_0^1 \left| DU(\lambda, r) - \alpha_0 - \alpha_1 W(r) \right|^2 dr \, .$$

We have the following theorem.

Theorem 3.1.A Let $\{y_t\}$ be generated under the null hypothesis (3.1) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} \ \mathrm{Wald}_\theta(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^*(\lambda, r)^2 dr \right)^{-1} \left(\int_0^1 DU^*(\lambda, r) dW(r) \right)^2,$$

$$\sup_{\lambda \in \Lambda} t_{|\theta|}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left| \left(\int_0^1 DU^*(\lambda, r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^*(\lambda, r) dW(r) \right) \right|,$$

$$\sup_{\lambda \in \Lambda} t_{\theta}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^*(\lambda, r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^*(\lambda, r) dW(r) \right),$$

and

$$\inf_{\lambda \in \Lambda} t_{\theta}(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^*(\lambda, r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^*(\lambda, r) dW(r) \right)$$

as $T \to \infty$, where $\Rightarrow$ denotes weak convergence in distribution in the sense of Billingsley (1968). The proof of this theorem proceeds along the lines of Zivot and Andrews (1992) and is therefore omitted. Perron's (1997) proof does not require that we consider the closed unit interval for λ. However, this only holds for unit root statistics, and not when a dummy variable statistic is used to determine the time of structural change.

We also consider the distribution of the step dummy t-statistic when the break date is chosen to minimize the unit root statistic. This is useful in circumstances where the unit root statistic is calculated as in Zivot and Andrews (1992), and then one wants to perform inference on θ. Following is the distribution of the Wald test for $\theta = 0$, choosing the break date to minimize t_ρ. We denote this statistic as $Wald_{\theta.\inf(\rho)}$.

Theorem 3.1.B Let $\{y_t\}$ be generated under the null hypothesis (3.1) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$Wald_{\theta.\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DU^*(\lambda^*,r)^2 dr\right)^{-1}\left(\int_0^1 DU^*(\lambda^*,r)dW(r)\right)^2$$

where

$$\lambda^* = \underset{\lambda \in \Lambda}{\arg\min} \left(\int_0^1 W^*(\lambda,r)^2 dr\right)^{-1/2}\left(\int_0^1 W^*(\lambda,r)dW(r)\right)$$

and the last term is the Perron and Vogelsang (1992) unit root statistic. Specifically, $W^*(\lambda,r)$ is the projection residual from the continuous time regression:

$$W(r) = \hat{\alpha}_0 + \hat{\alpha}_1 DU(\lambda,r) + W^*(\lambda,r) \ .$$

We now turn to the distribution of the Wald test of the null that $\theta = 0$ and $\rho = 1$. Let $\sup Wald_{\theta.\rho}$ denote this test statistic. Let $X_1(\lambda,r)' = (DU(\lambda,r), W(r))$ and $X_2(r) = 1$. Then $X^*_1(\lambda,r)$ is the projection residual from the continuous time regression which minimizes:

$$\min_{\alpha_0} \int_0^1 \left\| X_1(\lambda,r) - \alpha_0 X_2(r) \right\|^2 dr \ .$$

We then have the following result:

Theorem 3.1.C Let $\{y_t\}$ be generated under the null hypothesis (3.1) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of (0,1). Then,

$$\sup_{\lambda \in \Lambda} Wald_{\theta,\rho}(\lambda) \Rightarrow$$

$$\sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^*(\lambda,r)dW(r)\right)' \left(\int_0^1 X_1^*(\lambda,r)X_1^*(\lambda,r)'dr\right)^{-1} \left(\int_0^1 X_1^*(\lambda,r)dW(r)\right)$$

The asymptotic critical values for these statistics are presented in Table 3.5. The first row corresponds to the sup-Wald test analyzed by Vogelsang (1997) when the data are integrated. To simulate the asymptotic critical values, we set the sample size at 1000 and calculated the finite sample versions of the terms in Theorem 3.1 using Normal errors. We then repeated this 50,000 times. An upper bound on the standard errors of the critical values is 0.0022.

3.3.2 Trending data

We now turn to the analysis of trending data. Recall the null hypothesis and test regressions:

$$H_0 : y_t = \mu + y_{t-1} + u_t, \tag{3.4)'}$$

$$y_t = \hat{c} + \hat{\theta} DU(\hat{T}_B)_t + \hat{\beta} t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t, \tag{3.8)'}$$

$$y_t = \hat{c} + \hat{\gamma} DT(\hat{T}_B)_t + \hat{\beta} t + \hat{\rho} y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t, \tag{3.9)'}$$

and

$$y_t = \hat{c} + \hat{\theta}DU(\hat{T}_B)_t + \hat{\gamma}DT(\hat{T}_B)_t + \hat{\beta}t + \hat{\rho}y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \Delta y_{t-1} + \hat{e}_t \,. \qquad (3.10)'$$

For Model A we first consider $\sup W_\theta^A$, $\sup t_{\theta 1}^A$, $\sup t_\theta^A$, and $\inf t_\theta^A$. Likewise, for Model B we consider $\sup W_\gamma^B$, $\sup t_{|\gamma|}^B$, $\sup t_\gamma^B$, and $\inf t_\gamma^B$. Let $DU^A(\lambda,r)$ and $DT^B(\lambda,r)$ be the projection residuals from the following continuous time regressions:

$$DU(\lambda,r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + DU^A(\lambda,r) \,,$$

and

$$DT(\lambda,r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + DT^B(\lambda,r)$$

respectively.

Theorem 3.2.A. Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$\sup_{\lambda \in \Lambda} Wald_\theta^A(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^A(\lambda,r)^2 dr\right)^{-1} \left(\int_0^1 DU^A(\lambda,r) dW(r)\right)^2 ,$$

$$\sup_{\lambda \in \Lambda} \mathrm{t}^{A}_{\theta i}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left| \left(\int_0^1 DU^{A}(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^{A}(\lambda,r) dW(r) \right) \right|,$$

$$\sup_{\lambda \in \Lambda} t^{A}_{\theta}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^{A}(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^{A}(\lambda,r) dW(r) \right),$$

and

$$\inf_{\lambda \in \Lambda} t^{A}_{\theta}(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^{A}(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 DU^{A}(\lambda,r) dW(r) \right).$$

Similarly,

$$\sup_{\lambda \in \Lambda} Wald^{B}_{\lambda}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DT^{B}(\lambda,r)^2 dr \right)^{-1} \left(\int_0^1 DT^{B}(\lambda,r) dW(r) \right)^2,$$

$$\sup_{\lambda \in \Lambda} \mathrm{t}^{B}_{\lambda i}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left| \left(\int_0^1 DT^{B}(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 DT^{B}(\lambda,r) dW(r) \right) \right|,$$

$$\sup_{\lambda \in \Lambda} \mathrm{t}^{B}_{\lambda}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DT^{B}(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 DT^{B}(\lambda,r) dW(r) \right),$$

and

$$\inf_{\lambda \in \Lambda} t_{\lambda}^{B}(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DT^B(\lambda,r)^2 dr\right)^{-1/2} \left(\int_0^1 DT^B(\lambda,r) dW(r)\right).$$

Critical values for these statistics are presented in Tables 3.6 and 3.7 respectively. The first row in each table corresponds to the F-statistic computed by Banerjee et. al (1992). Also reported in the 5th rows of Tables 3.6 and 3.7 are $\sup Wald^{A}_{\theta,\inf(\rho)}$ and $\sup Wald^{B}_{\gamma,\inf(\rho)}$, the Wald tests of no structural change when the break is chosen to minimize the unit root statistic. These distributions are derived in the following theorem.

Theorem 3.2.B. Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$Wald^{A}_{\theta,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DU^A(\lambda^A,r)^2 dr\right)^{-1} \left(\int_0^1 DU^A(\lambda^A,r) dW(r)\right)^2$$

and

$$Wald^{B}_{\gamma,\inf(\rho)}(\lambda) \Rightarrow \left(\int_0^1 DT^B(\lambda^B,r)^2 dr\right)^{-1} \left(\int_0^1 DT^B(\lambda^B,r) dW(r)\right)^2$$

where

$$\lambda^A = \arg\min_{\lambda \in \Lambda} \left(\int_0^1 W^A(\lambda,r)^2 dr\right)^{-1/2} \left(\int_0^1 W^A(\lambda,r) dW(r)\right)$$

and

$$\lambda^B = \underset{\lambda \in \Lambda}{\arg\min} \left(\int_0^1 W^B(\lambda,r)^2 dr \right)^{-1/2} \left(\int_0^1 W^B(\lambda,r) dW(r) \right),$$

where the last two terms are the Zivot-Andrews unit root tests for Models A and B respectively. That is, $W^A(\lambda,r)$ and $W^B(\lambda,r)$ are the projection residuals from the continuous time regressions

$$W(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 DU(\lambda,r) + W^A(\lambda,r)$$

and

$$W(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 DT(\lambda,r) + W^B(\lambda,r)$$

respectively.

Finally for Models A and B, we derive the limiting distributions for $\sup Wald^A_{\theta,\rho}$ and $\sup Wald^B_{\gamma,\rho}$, the tests of the joint null hypothesis of a unit root and no structural change. Let $X_1^A(\lambda,r)' = (DU(\lambda,r), W(r))$, and $X_1^B(\lambda,r)' = (DT(\lambda,r), W(r))$. We then have the following.

Theorem 3.2.C Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$\sup_{\lambda \in \Lambda} Wald^A_{\theta,\rho}(\lambda) \Rightarrow$$

$$\sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{A*}(\lambda,r) dW(r) \right)' \left(\int_0^1 X_1^{A*}(\lambda,r) X_1^{A*}(\lambda,r)' dr \right)^{-1} \left(\int_0^1 X_1^{A*}(\lambda,r) dW(r) \right)$$

and

$$\sup_{\lambda \in \Lambda} Wald_{\gamma,\rho}^{B}(\lambda) \Rightarrow$$

$$\sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{B*}(\lambda,r) dW(r) \right)' \left(\int_0^1 X_1^{B*}(\lambda,r) X_1^{B*}(\lambda,r)' dr \right)^{-1} \left(\int_0^1 X_1^{B*}(\lambda,r) dW(r) \right).$$

Critical values for these are reported in the 6th column of Tables 3.6 and 3.7.

We now turn to Model C. For this model, we consider 8 individual test statistics for θ and γ. These are $\inf t_\theta^C$, $\sup t_\theta^C$, $\sup t_{|\theta|}^C$, $\sup Wald_\theta^C$, $\inf t_\gamma^C$, $\sup t_\gamma^C$, $\sup t_{|\gamma|}^C$, and $\sup Wald_\gamma^C$. Let $DU^C(\lambda,r)$ and $DT^C(\lambda,r)$ be the projection residuals from the following continuous time regressions:

$$DU(\lambda,r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + \hat{\alpha}_3 DT(\lambda,r) + DU^C(\lambda,r)$$

and

$$DT(\lambda,r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 W(r) + \hat{\alpha}_3 DU(\lambda,r) + DT^C(\lambda,r)$$

respectively.

The distribution of these statistics are presented in the following theorem.

Theorem 3.2.D Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$\sup_{\lambda \in \Lambda} Wald_{\theta}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^{C}(\lambda,r)^2\,dr\right)^{-1}\left(\int_0^1 DU^{C}(\lambda,r)\,dW(r)\right)^2,$$

$$\sup_{\lambda \in \Lambda} t_{|\theta|}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left|\left(\int_0^1 DU^{C}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 DU^{C}(\lambda,r)\,dW(r)\right)\right|,$$

$$\sup_{\lambda \in \Lambda} t_{\theta}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DU^{C}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 DU^{C}(\lambda,r)\,dW(r)\right),$$

$$\inf_{\lambda \in \Lambda} t_{\theta}^{C}(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DU^{C}(\lambda,r)^2\,dr\right)^{-1/2}\left(\int_0^1 DU^{C}(\lambda,r)\,dW(r)\right),$$

$$\sup_{\lambda \in \Lambda} Wald_{\gamma}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DT^{C}(\lambda,r)^2\,dr\right)^{-1}\left(\int_0^1 DT^{C}(\lambda,r)\,dW(r)\right)^2,$$

$$\sup_{\lambda \in \Lambda} t_{\gamma 1}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left| \left(\int_0^1 DT^C(\lambda,r)^2 dr\right)^{-1/2} \left(\int_0^1 DT^C(\lambda,r) dW(r)\right) \right|,$$

$$\sup_{\lambda \in \Lambda} t_{\gamma}^{C}(\lambda) \Rightarrow \sup_{\lambda \in \Lambda} \left(\int_0^1 DT^C(\lambda,r)^2 dr\right)^{-1/2} \left(\int_0^1 DT^C(\lambda,r) dW(r)\right),$$

and

$$\inf_{\lambda \in \Lambda} t_{\gamma}^{C}(\lambda) \Rightarrow \inf_{\lambda \in \Lambda} \left(\int_0^1 DT^C(\lambda,r)^2 dr\right)^{-1/2} \left(\int_0^1 DT^C(\lambda,r) dW(r)\right).$$

Asymptotic critical values for these statistics appear in the first 8 rows of Table 3.8.

We next consider the Wald test for the joint hypothesis that there is neither a level shift nor a change in growth $(\theta = \gamma = 0)$. Let $X_1^C(\lambda,r)' = (DU(\lambda,r), DT(\lambda,r))$. We then have the following result.

Theorem 3.2.E Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$\sup_{\lambda \in \Lambda} Wald_{\theta,\gamma}^{C}(\lambda) \Rightarrow$$

$$\sup_{\lambda \in \Lambda} \left(\int_0^1 X_1^{C*}(\lambda,r)dW(r)\right)'\left(\int_0^1 X_1^{C*}(\lambda,r)X_1^{C*}(\lambda,r)'dr\right)^{-1}\left(\int_0^1 X_1^{C*}(\lambda,r)dW(r)\right).$$

Critical values for this statistic are in the 9th row of Table 3.8.

What is also potentially of interest is the Wald test of no structural change when λ minimizes the unit root statistic. We denote this statistic as $Wald^{C}_{\theta,\gamma,\inf(\rho)}$. This distribution is presented in the following theorem.

Theorem 3.2.F Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$Wald^{C}_{\theta,\gamma,\inf(\rho)}(\lambda) \Rightarrow$$

$$\left(\int_0^1 X_1^{C*}(\lambda^C,r)dW(r)\right)'\left(\int_0^1 X_1^{C*}(\lambda^C,r)X_1^{C*}(\lambda^C,r)'dr\right)^{-1}\left(\int_0^1 X_1^{C*}(\lambda^C,r)dW(r)\right)$$

where

$$\lambda^C = \underset{\lambda \in \Lambda}{\arg\min} \left(\int_0^1 W^C(\lambda,r)^2 dr\right)^{-1/2}\left(\int_0^1 W^C(\lambda,r)dW(r)\right)$$

is the Zivot Andrews Model C unit root statistic, *i.e.* $W^C(\lambda,r)$ is the projection residual from the continuous time regression:

$$W(r) = \hat{\alpha}_0 + \hat{\alpha}_1 r + \hat{\alpha}_2 DU(\lambda,r) + \hat{\alpha}_C DT(\lambda,r) + W^C(\lambda,r) \ .$$

Critical values for this statistic are in the 10th row of Table 3.8.

We conclude this section with the joint distribution of the Wald test that there is a unit root and no structural change $(\rho = 1, \theta = \gamma = 0)$. Denote this statistic by $\sup Wald^{C}_{\theta,\gamma,\rho}$. Let $X^{C}_{1}(\lambda, r)' = (DU(\lambda, r), DT(\lambda, r), W(r))$. We then have,

Theorem 3.2.G Let $\{y_t\}$ be generated under the null hypothesis (3.4) and let $\{u_t\}$ be i.i.d., mean 0, with $0 < \sigma^2 < \infty$. Let Λ be a closed subset of $(0,1)$. Then,

$$\sup_{\lambda \in \Lambda} Wald^{C}_{\theta,\gamma,\rho}(\lambda) \Rightarrow$$

$$\sup_{\lambda \in \Lambda} \left(\int_0^1 X^{C*}_{1}(\lambda^C, r)\,dW(r)\right)' \left(\int_0^1 X^{C*}_{1}(\lambda^C, r) X^{C*}_{1}(\lambda^C, r)'\,dr\right)^{-1} \left(\int_0^1 X^{C*}_{1}(\lambda^C, r)\,dW(r)\right).$$

The last row of Table 3.8 presents the asymptotic critical values for this statistic.

3.4. *Finite Sample Size and Power*

In this section, we ascertain the finite sample properties of the statistics presented in Section 3.3, in terms of size and power. Table 3.9 presents the empirical size of selected tests statistics at the nominal 0.05 and 0.10 significance levels for samples sizes T=100 and 200. We set the number of iterations to 5,000 for all simulations to follow.

Generally, the asymptotic critical values provide a reasonable approximation to the finite sample distributions, for samples sizes as small at 100. Doubling the samples size only results in a slight mitigation of the size distortion.

We not turn to the relative power of the joint and individual test statistics. We consider two values for ρ under the alternative hypothesis; 0.9 and 0.7. The range of θ and γ considered are 0.5, 1, and 2. The last value corresponds to a trend break 2 times the size of the innovation standard deviation. We also consider negative values of θ and γ. The results for negative values are not qualitatively different from the positive values and are available upon request from the author.

Table 3.10 presents the size-adjusted power for non-trending data at the 0.05 and 0.10 significance levels. A few points are in order. First, the 1-sided test for no structural change has higher power than the 2-sided test. This result corroborates the finding of Vogelsang and Perron (1994) and Perron (1997) that imposing a sign for the trend break leads to an increase in power. Second, the $\sup Wald_{\theta,\rho}$ statistic uniformly dominates the $\sup Wald_{\theta}$ and $\sup t_{\theta}$ for all values of θ and ρ considered. However, all statistics are dominated by the 1-sided unit root test. This result is analogous to a finding of Dickey, Bell, and Miller (1986). They demonstrate that in standard ADF tests, the 1-sided t-statistic for the unit root null has more power than the 2-sided F-test of the joint null of a unit root and no time trend.

Thus, in practice, if the researcher is only interested on whether or not the time series has a unit root, the $\inf t_{\rho}$ should be used. However, if the researcher is interested in performing inference on θ, then for non-trending data, the $\sup Wald_{\theta,\rho}$ should be implemented.

A different picture emerges for trending data. Tables 3.11-3.13 present size adjusted power for Models A, B, and C respectively. To conserve space, we only present results for θ and γ both equal to 0.5 or 1.0 for Model C. For Model A, the test which has the highest power depends on ρ. For ρ=0.9 and a level shift in the range of 1 to 2 innovation standard deviations, the $\sup t_{\theta}^{A}$ test is the clear winner. However, for the less persistent alternative (ρ=0.7), the $\sup Wald_{\theta,\rho}^{A}$ test outperforms the $\sup t_{\theta}^{A}$ test.

Turning now to Model B, for changes in growth in the range of 0.5 to 1 innovation standard deviation, the 1-sided test, $\sup t_{\gamma}^{B}$, is the clear winner for both values of T and ρ. Furthermore, the 1-sided unit root test, $\inf t_{\rho}^{B}$, dominates the joint test. There does not appear to be a distinct advantage to performing the $\sup Wald_{\gamma,\rho}^{B}$ test over the 1-sided tests for structural change or a unit root. For larger changes in growth, all tests perform remarkably well. But as we shall see in the next section, changes in growth this size do not occur in U.S. output. Therefore we do not report power results for trend breaks greater than 2 innovation standard deviations.

A similar pictures emerges for Model C. The joint Wald test of a unit root and no structural change, $\sup Wald_{\theta,\gamma,\rho}^{C}$ is dominated by every individual 1-sided test with the exception of $\sup t_{\theta}^{C}$. Also of interest is that the joint test that θ=γ=0 has less power than the 2-sided test that γ=0.

3.5 Application to U. S. GDP

As an empirical application, we reconsider the Zivot-Andrews unit root tests on annual and quarterly U.S. real GDP analyzed by Murray and Nelson (1998). Murray and Nelson perform the Model A unit root test on the Maddison (1995) annual GDP series (1870-1994), and the Model B unit root test on post-war quarterly chained U.S. GDP (1947.1–1997.3). They demonstrate that whether the lag length is selected by the general to specific (GS) strategy or the Schwartz information criterion (SIC), the unit root null is rejected at the 0.05 level for annual GDP, but not at the 0.10 level for quarterly GDP. These regressions are presented in Table 3.14. Since each test considered in this section chooses the same break date for annual and quarterly data (1929 and 1972.2), we present the results for each series and lag selection procedure as one regression. As in Perron (1989) and Zivot and Andrews (1992), the maximum lag length considered is 8 for annual data and 12 for quarterly data. For either frequency, GS chooses the maximum lag allowed, while SIC chooses only 1.

Using the critical values from the 5th row in Tables 3.6 and 3.7, we can assess whether or not $\hat{\theta}$ or $\hat{\gamma}$, the step and ramp dummy coefficients, are significant when the break date is chosen to minimize t_{ρ}. (These are the $Wald^{A}_{\theta,\inf(\rho)}$ and $Wald^{B}_{\gamma,\inf(\rho)}$ statistics). For the annual series (Model A), the level shift is significant at the 0.10 level for GS, but insignificant for SIC. While both methods of lag selection result in rejection of the unit root null, GS suggests stationarity around a broken trend, while SIC indicates stationarity around a constant trend.

To assess whether there has been structural change while not explicitly testing the unit root hypothesis, we perform the 1-sided $\inf t_{\theta}^{A}$ test. This statistic is significant at the 0.05 level for GS, but not significant at the 0.10 level for SIC.

We now turn our attention to the $\sup Wald_{\theta.\rho}^{A}$ statistic, which tests the joint null hypothesis. This statistic is significant at the 0.05 level for both methods of lag selection. It thus suggests that GDP is stationary around a broken trend. The disagreement between $\sup Wald_{\theta.\rho}^{A}$ and $\inf t_{\theta}^{A}$, for SIC, may be due to poor power properties of $\inf t_{\theta}^{A}$ when only a subset of the null is violated, *i.e.* $\rho < 1$ and $\theta \neq 0$.

In Section 3.4 we demonstrated that for level shifts of the size estimated for this series (1 to 2 innovation standard deviations), and a non-local autoregressive root (0.7), the power of the $\sup Wald_{\theta.\rho}^{A}$ statistic dominates the 1-sided tests for structural change, but is dominated by the 1-sided Zivot-Andrews unit root test. Given that none of the statistics have an appreciable finite sample size distortion, these results lead us to conclude that annual GDP is stationary around a broken trend.

Analogous statistics for quarterly GDP are also presented in Table 3.14. For this series, which appears to have a unit root, we find that the pre and post break growth rates, based on the $Wald_{\gamma,\inf(\rho)}^{B}$ statistic, are not statistically different under either method of lag selection. A different picture emerges if we compute the 1-sided $\inf t_{\gamma}^{B}$

test for a change in growth. Under GS, there is not a statistical difference in growth rates, but the SIC results in a rejection of the null at the 0.10 level.

Turning to the $\sup Wald^{B}_{\gamma,\rho}$, for both methods of lag selection the joint test corroborates the Zivot-Andrews unit root test. Neither is significant at the 0.10 level.

We demonstrated in Section 4 that for the small changes in growth (less that 1 innovation standard deviation) that this series appears to exhibit, the 1-sided tests for structural change uniformly dominate all other statistics in terms of power. Since both methods of lag selection lead to different outcomes for the $\inf t^{B}_{\gamma}$ test, we can conclude that there is a unit root, but we are uncertain as to whether the rate of growth has changed in the postwar period.

3.6 Conclusion

The purpose of this paper has been to fill in the gaps in the literature concerning the asymptotic distributions of test statistics for a unit root and/or structural change. We derive 1 and 2-sided tests for the null of no structural change as well as joint tests of the hypothesis that a time series is integrated without structural change. The motivation for the latter is the potential increase in power over tests which do not explicitly test the unit root hypothesis.

For Model A, no clear winner emerges. For level shifts of the size estimated for U.S. real GDP, the joint test has higher power than individual tests for non-local autoregressive roots. However, the situation is reversed for a local root. For Model

B, the 1-sided tests for structural change dominate the joint tests for small changes in growth, regardless of the size of the autoregressive root.

We apply the tests derived here to annual and quarterly U.S. real GDP. Almost all tests agree that the 1870-1994 annual GDP series is stationary around a broken time trend with a change in level occurring at 1929. While all tests indicate that the 1947.1-1997.3 quarterly GDP series has a unit root, there is not a consensus as to whether or not the growth rate began to slow in 1972.2.

Table 3.1 Non-trending Data in the Literature

Statistic	Origin
$\inf t_\rho$	PV: AO/IO
$t_{\rho,\inf(\theta)}$	PV: AO/IO
$Wald_{\theta,\inf(\rho)}$	
$\sup Wald_\theta$	V: AO - P: AI/IO
$\sup Wald_{\theta,\rho}$	

PV is Perron and Vogelsang (1992). V is Vogelsang (1997)

Table 3.2 Trending Data in the Literature; Model A

Statistic	Origin
$\inf t_\rho^A$	ZA: IO - BLS: IO VP: AO - P: IO
$t^A_{\rho,Wald(\theta)}$	BLS: IO VP: AO - P: IO
$Wald^A_{\theta,\inf(\rho)}$	
$\sup Wald_\theta^A$	BLS: IO
$\sup Wald^A_{\theta,\rho}$	

ZA is Zivot and Andrews (1992). BLS is Banerjee, Lumsdaine, and Stock (1992). VP is Vogelsang and Perron (1994). P is Perron (1997).

Table 3.3 Trending Data in the Literature; Model B

Statistic	Origin
$\inf t_{\rho}^{B}$	ZA: IO - BLS: IO VP: AO - P: AO
$t_{\rho,Wald(\gamma)}^{B}$	BLS: IO VP: AO - P: AO
$Wald_{\gamma,\inf(\rho)}^{B}$	
$\sup Wald_{\gamma}^{B}$	BLS: IO
$Wald_{\gamma,\inf(\rho)}^{B}$	

ZA is Zivot and Andrews (1992). BLS is Banerjee, Lumsdaine, and Stock (1992). VP is Vogelsang and Perron (1994). P is Perron (1997).

Table 3.4 Trending Data in the Literature; Model C

Statistic	Origin
$\inf t_{\rho}^{C}$	ZA: IO VP: AO - P: IO
$t_{\rho,Wald(\gamma)}^{C}$	VP: AO - P: AO
$t_{\rho,Wald(\theta,\rho)}^{C}$	
$Wald_{\theta,\gamma,\inf(\rho)}^{C}$	
$\sup Wald_{\theta}^{C}$	
$\sup Wald_{\gamma}^{C}$	
$\sup Wald_{\theta,\gamma}^{C}$	V: AO
$\sup Wald_{\theta,\gamma,\rho}^{C}$	

Table 3.5 Asymptotic Critical Values; Non-trending Data

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
sup $Wald_\theta$	22.477	19.934	17.864	15.633	9.538	5.501	4.683	4.085	3.524
sup $t_{\theta l}$	4.741	4.465	4.227	3.954	3.088	2.346	2.164	2.021	1.877
sup t_θ	4.558	4.229	3.959	3.650	2.370	1.102	0.849	0.634	0.393
inf t_θ	-4.519	-4.222	-3.950	-3.643	-2.352	-1.102	-0.836	-0.621	-0.400
$Wald_{\theta,\inf(\rho)}$	21.975	19.464	17.442	15.236	8.900	4.367	3.441	2.809	2.276
sup $Wald_{\theta,\rho}$	25.305	22.798	20.742	18.610	12.625	8.638	7.7979	7.126	6.374

Table 3.6 Asymptotic Critical Values; Model A

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
sup $Wald_\theta$	24.130	21.547	19.448	17.315	11.371	7.181	6.293	5.606	4.906
sup $t_{\theta l}$	4.912	4.642	4.410	4.161	3.372	2.680	2.509	2.368	2.215
sup t_θ	4.705	4.404	4.154	3.871	2.971	2.244	2.077	1.948	1.805
inf t_θ	-4.723	-4.424	-4.174	-3.896	-2.975	-2.239	-2.074	-1.947	-1.811
$Wald_{\theta,\inf(\rho)}$	23.870	21.179	19.067	16.830	10.649	6.231	5.280	4.598	3.929
sup $Wald_{\theta,\rho}$	29.843	27.195	25.141	22.898	16.618	12.346	11.390	10.534	9.587

Table 3.7 Asymptotic Critical Values; Model B

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
sup $Wald_\gamma$	20.710	17.831	15.697	13.261	6.502	2.828	2.221	1.808	1.411
sup $t_{\gamma l}$	4.551	4.223	3.962	3.642	2.550	1.682	1.490	1.345	1.188
sup t_γ	4.321	3.968	3.650	3.261	1.787	0.576	0.292	0.058	-0.212
inf t_γ	-4.291	-3.953	-3.634	-3.269	-1.783	-0.579	-0.293	-0.052	0.216
$Wald_{\gamma,\inf(\rho)}$	20.589	17.814	15.521	13.101	5.683	1.889	1.343	0.996	0.701
sup $Wald_{\gamma,\rho}$	25.168	22.635	20.542	18.363	12.114	7.908	7.023	6.398	5.768

Table 3.8 Asymptotic Critical Values; Model C

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
sup $Wald_{\theta}$	24.852	22.143	20.033	17.878	11.859	8.124	7.307	6.676	6.011
sup $t_{\lvert\theta\rvert}$	4.985	4.706	4.476	4.228	3.444	2.850	2.703	2.584	2.452
sup t_{θ}	4.785	4.469	4.221	3.935	3.035	2.339	2.179	2.047	1.901
inf t_{θ}	-4.776	-4.485	-4.236	-3.939	-3.040	-2.346	-2.182	-2.056	-1.919
sup $Wald_{\gamma}$	25.044	21.889	19.355	16.689	9.597	5.739	4.995	4.467	3.932
sup $t_{\lvert\gamma\rvert}$	5.004	4.679	4.400	4.085	3.098	2.396	2.235	2.114	1.983
sup t_{γ}	4.781	4.397	4.086	3.721	2.548	1.670	1.458	1.283	1.092
inf t_{γ}	-4.749	-4.404	-4.086	-3.717	-2.544	-1.669	-1.457	-1.285	-1.091
sup $Wald_{\theta,\gamma}$	29.296	26.423	24.095	21.695	14.734	10.008	9.011	8.223	7.386
$Wald_{\theta,\gamma,inf(\rho)}$	28.807	26.071	23.683	21.126	13.592	8.140	6.969	6.093	5.213
sup $Wald_{\theta,\gamma,\rho}$	34.244	31.430	29.264	26.881	20.100	15.436	14.408	13.552	12.603

Table 3.9 Finite Sample Size

	T=100		T=200	
	5.0%	10.0%	5.0%	10.0%
Non-trending data				
sup $Wald_{\theta}$	0.047	0.096	0.051	0.102
sup $t_{\vert\theta\vert}$	0.048	0.096	0.051	0.102
sup t_{θ}	0.048	0.086	0.047	0.092
inf t_{θ}	0.047	0.086	0.054	0.101
sup $Wald_{\theta,\rho}$	0.054	0.101	0.052	0.098
Model A				
sup $Wald_{\theta}$	0.041	0.076	0.442	0.087
sup $t_{\vert\theta\vert}$	0.041	0.076	0.442	0.087
sup t_{θ}	0.040	0.075	0.502	0.094
inf t_{θ}	0.038	0.074	0.380	0.078
sup $Wald_{\theta,\rho}$	0.059	0.108	0.056	0.104
Model B				
sup $Wald_{\gamma}$	0.065	0.116	0.055	0.109
sup $t_{\vert\gamma\vert}$	0.065	0.116	0.055	0.109
sup t_{γ}	0.055	0.104	0.057	0.101
inf t_{γ}	0.062	0.106	0.057	0.101
sup $Wald_{\gamma,\rho}$	0.075	0.138	0.062	0.120
Model C				
sup $Wald_{\theta}$	0.040	0.074	0.048	0.090
sup $t_{\vert\theta\vert}$	0.040	0.074	0.048	0.090
sup t_{θ}	0.038	0.074	0.045	0.086
inf t_{θ}	0.036	0.071	0.045	0.089
sup $Wald_{\gamma}$	0.058	0.104	0.046	0.098
sup $t_{\vert\gamma\vert}$	0.058	0.104	0.046	0.098
sup t_{γ}	0.049	0.089	0.049	0.095
inf t_{γ}	0.055	0.096	0.049	0.094
sup $Wald_{\theta,\gamma}$	0.057	0.100	0.053	0.099
sup $Wald_{\theta,\gamma,\rho}$	0.660	0.117	0.592	0.109

Table 3.10 Size Adjusted Power; Non-trending Data

			T=100		T=200	
			5.0%	10.0%	5.0%	10.0%
ρ=0.9						
	θ=0.5	sup Wald$_\theta$	0.020	0.041	0.008	0.020
		sup t_θ	0.014	0.028	0.015	0.036
		inf t_ρ	0.134	0.228	0.391	0.572
		sup Wald$_{\theta,\rho}$	0.105	0.191	0.337	0.516
	θ=1.0	sup Wald$_\theta$	0.021	0.042	0.013	0.028
		sup t_θ	0.035	0.080	0.026	0.057
		inf t_ρ	0.142	0.234	0.389	0.570
		sup Wald$_{\theta,\rho}$	0.100	0.194	0.336	0.515
	θ=2.0	sup Wald$_\theta$	0.036	0.067	0.022	0.048
		sup t_θ	0.064	0.137	0.050	0.107
		inf t_ρ	0.146	0.250	0.382	0.562
		sup Wald$_{\theta,\rho}$	0.107	0.196	0.323	0.494
ρ=0.7						
	θ=0.5	sup Wald$_\theta$	0.011	0.024	0.011	0.019
		sup t_θ	0.021	0.050	0.018	0.044
		inf t_ρ	0.838	0.923	1.000	1.000
		sup Wald$_{\theta,\rho}$	0.758	0.885	1.000	1.000
	θ=1.0	sup Wald$_\theta$	0.025	0.056	0.034	0.067
		sup t_θ	0.054	0.113	0.069	0.129
		inf t_ρ	0.842	0.931	1.000	1.000
		sup Wald$_{\theta,\rho}$	0.756	0.881	1.000	1.000
	θ=2.0	sup Wald$_\theta$	0.137	0.236	0.284	0.418
		sup t_θ	0.237	0.385	0.430	0.593
		inf t_ρ	0.842	0.928	1.000	1.000
		sup Wald$_{\theta,\rho}$	0.735	0.865	1.000	1.000

Table 3.11 Size Adjusted Power; Model A

		T=100		T=200	
		5.0%	10.0%	5.0%	10.0%
ρ=0.9					
θ=0.5	sup $Wald_\theta$	0.042	0.085	0.052	0.099
	sup t_θ	0.057	0.112	0.035	0.070
	inf t_ρ	0.096	0.188	0.271	0.430
	sup $Wald_{\theta,\rho}$	0.090	0.161	0.213	0.359
θ=1.0	sup $Wald_\theta$	0.116	0.198	0.226	0.344
	sup t_θ	0.183	0.301	0.318	0.465
	inf t_ρ	0.081	0.149	0.299	0.466
	sup $Wald_{\theta,\rho}$	0.095	0.169	0.282	0.443
θ=2.0	sup $Wald_\theta$	0.732	0.831	0.947	0.976
	sup t_θ	0.827	0.911	0.972	0.991
	inf t_ρ	0.243	0.386	0.791	0.895
	sup $Wald_{\theta,\rho}$	0.582	0.706	0.892	0.946
ρ=0.7					
θ=0.5	sup $Wald_\theta$	0.019	0.039	0.021	0.040
	sup t_θ	0.027	0.054	0.030	0.067
	inf t_ρ	0.707	0.850	1.000	1.000
	sup $Wald_{\theta,\rho}$	0.647	0.789	0.998	1.000
θ=1.0	sup $Wald_\theta$	0.066	0.124	0.205	0.306
	sup t_θ	0.116	0.205	0.286	0.422
	inf t_ρ	0.639	0.811	0.999	0.999
	sup $Wald_{\theta,\rho}$	0.582	0.733	0.998	0.999
θ=2.0	sup $Wald_\theta$	0.548	0.694	0.949	0.978
	sup t_θ	0.690	0.815	0.975	0.978
	inf t_ρ	0.751	0.877	1.000	1.000
	sup $Wald_{\theta,\rho}$	0.743	0.844	0.996	1.000

Table 3.12 Size Adjusted Power; Model B

			T=100		T=200	
			5.0%	10.0%	5.0%	10.0%
ρ=0.9						
	γ=0.5	sup $Wald_\gamma$	0.192	0.332	0.612	0.793
		sup t_γ	0.348	0.586	0.799	0.938
		inf t_ρ	0.109	0.214	0.409	0.609
		sup $Wald_{\gamma,\rho}$	0.095	0.179	0.336	0.506
	γ=1.0	sup $Wald_\gamma$	0.513	0.691	0.973	0.994
		sup t_γ	0.705	0.878	0.994	0.998
		inf t_ρ	0.274	0.454	0.893	0.956
		sup $Wald_{\gamma,\rho}$	0.241	0.395	0.850	0.928
	γ=2.0	sup $Wald_\gamma$	0.979	0.994	1.000	1.000
		sup t_γ	0.995	0.999	1.000	1.000
		inf t_ρ	0.930	0.966	1.000	1.000
		sup $Wald_{\gamma,\rho}$	0.920	0.960	1.000	1.000
	γ=3.0	sup $Wald_\gamma$	1.000	1.000	1.000	1.000
		sup t_γ	1.000	1.000	1.000	1.000
		inf t_ρ	0.999	0.999	1.000	1.000
		sup $Wald_{\gamma,\rho}$	0.999	1.000	1.000	1.000
ρ=0.7						
	γ=0.5	sup $Wald_\gamma$	0.814	0.926	1.000	1.000
		sup t_γ	0.932	0.989	1.000	1.000
		inf t_ρ	0.735	0.883	1.000	1.000
		sup $Wald_{\gamma,\rho}$	0.687	0.835	0.999	1.000
	γ=1.0	sup $Wald_\gamma$	0.976	0.995	1.000	1.000
		sup t_γ	0.996	0.999	1.000	1.000
		inf t_ρ	0.902	0.967	1.000	1.000
		sup $Wald_{\gamma,\rho}$	0.868	0.950	1.000	1.000
	γ=2.0	sup $Wald_\gamma$	0.999	1.000	1.000	1.000
		sup t_γ	1.000	1.000	1.000	1.000
		inf t_ρ	0.999	0.999	1.000	1.000
		sup $Wald_{\gamma,\rho}$	0.999	0.999	1.000	1.000

Table 3.13 Size Adjusted Power; Model C

		T=100		T=200	
		5.0%	10.0%	5.0%	10.0%
ρ=0.9					
θ=0.5	sup $Wald_\theta$	0.068	0.127	0.115	0.208
γ=0.5	sup t_θ	0.091	0.164	0.183	0.289
	sup $Wald_\gamma$	0.152	0.274	0.544	0.705
	sup t_γ	0.294	0.489	0.707	0.874
	inf t_ρ	0.079	0.150	0.221	0.370
	sup $Wald_{\theta,\gamma}$	0.116	0.213	0.348	0.518
	sup $Wald_{\theta,\gamma,\rho}$	0.059	0.117	0.167	0.283
θ=1.0	sup $Wald_\theta$	0.226	0.323	0.557	0.686
γ=1.0	sup t_θ	0.289	0.399	0.681	0.794
	sup $Wald_\gamma$	0.396	0.580	0.905	0.965
	sup t_γ	0.602	0.783	0.966	0.993
	inf t_ρ	0.134	0.243	0.493	0.678
	sup $Wald_{\theta,\gamma}$	0.356	0.512	0.857	0.932
	sup $Wald_{\theta,\gamma,\rho}$	0.169	0.293	0.621	0.763
ρ=0.7					
θ=0.5	sup $Wald_\theta$	0.084	0.161	0.524	0.694
γ=0.5	sup t_θ	0.111	0.205	0.621	0.783
	sup $Wald_\gamma$	0.689	0.853	1.000	1.000
	sup t_γ	0.867	0.958	1.000	1.000
	inf t_ρ	0.557	0.728	0.998	1.000
	sup $Wald_{\theta,\gamma}$	0.529	0.709	0.999	1.000
	sup $Wald_{\theta,\gamma,\rho}$	0.404	0.583	0.992	0.998
θ=1.0	sup $Wald_\theta$	0.283	0.429	0.971	0.992
γ=1.0	sup t_θ	0.391	0.540	0.992	0.999
	sup $Wald_\gamma$	0.921	0.979	1.000	1.000
	sup t_γ	0.980	0.999	1.000	1.000
	inf t_ρ	0.658	0.817	1.000	1.000
	sup $Wald_{\theta,\gamma}$	0.842	0.930	1.000	1.000
	sup $Wald_{\theta,\gamma,\rho}$	0.617	0.787	1.000	1.000

Table 3.14 Application to U.S. GDP

Model A: Annual U.S. GDP 1870-1994								
	T_B	k	c	θ	β	ρ	$W_{\theta,\inf(\rho)}$	$\sup W_{\theta,\rho}$
GS	1929	8	1.574	-0.103	0.018	0.469	18.775*	37.371***
			(6.268)	(-4.333)**	(6.071)	(-6.105)***		
SIC	1929	1	0.787	-0.059	0.009	0.744	7.935	26.057**
			(5.263)	(-2.817)	(5.047)	(-5.010)**		

Model B: Quarterly U.S. GDP 1947.1 - 1997.3								
	T_B	k	c	γ	β	ρ	$W_{\gamma,\inf(\rho)}$	$\mathrm{Sup}W_{\gamma,\rho}$
GS	1972.2	12	0.828	-0.0003	0.001	0.886	8.567	10.979
			(3.359)	(-2.927)	(3.3216)	(-3.313)		
SIC	1972.2	1	0.720	-0.0003	0.001	0.901	11.498	15.950
			(4.030)	(-3.391)*	(3.920)	(-3.992)		

t-statistics are in parentheses
*,**,and *** denotes significance at the 10%, 5%, and 1% significance levels respectively

BIBLIOGRAPHY

Andrews, D. W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59:817-858.

Andrews, D.W.K., 1993, Tests for parameter instability and structural change with unknown change point, *Econometrica*, 61:821-56.

Andrews, D.W.K., and Werner Ploberger, 1994, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 62:1383-1414.

Baily, Martin Neil, and Robert J. Gordon, 1988, The productivity slowdown, measurement issues, and the explosion of computer power, *Brookings Papers on Economic Activity*, 0:347-420.

Balke, Nathan S., and Thomas B. Fomby, 1991, Shifting trends, segmented trends, and infrequent permanent shocks, *Journal of Monetary Economics*, 28:61-86.

Balke, Nathan S., and Robert J. Gordon, 1989, The prewar business cycle reconsidered: New estimates of gross national product 1869-1908, *Journal of Political Economy*, 97:38-92.

Banerjee, Anindya, Robin L. Lumsdaine, and James H. Stock, 1992, Recursive and sequential tests of the unit-toot and trend-break hypotheses: Theory and international evidence, *Journal of Business and Economic Statistics*, 10:271-287.

Beaudry, P. and G. Koop, 1993, Do recessions permanently change output?, *Journal of Monetary Economics*, 31:149-163.

Ben-David, Dan, and David H. Papell, 1995, The great wars, the great crash, and steady state growth: Some new evidence about an old stylized fact, *Journal of Monetary Economics*, 36:453-75.

Beveridge, Stephen and Charles R. Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of Monetary Economics*, 7:151-174.

Billingsley, Patrick, 1968, *Convergence of probability measures*, Wiley, New York.

Box, George E. P. and Gwilym, M. Jenkins, 1970, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

Brock, W.A. and C. L. Sayers, 1988, Is the business cycle characterized by deterministic chaos?, *Journal of Monetary Economics*, 22:71-90

Burns, A. F. and W. A. Mitchell, 1946, *Measuring business cycles*, NBER, New York, NY.

Campbell, John Y. and Pierre Perron, 1991, Pitfalls and opportunities: What macroeconomists should know about unit roots, in *NBER Macroeconomic Annual*, eds. O. J. Blanchard and S. Fischer, MIT Press, Cambridge, 141-201.

Chauvet, Marcelle, 1997, An econometric characterization of business cycle dynamics with factor structure and regime switching, forthcoming in *International Economic Review.*

Cheung, Yin-Wong, and Menzie D. Chinn, 1997, Further investigation of the uncertain unit root in GNP, *Journal of Business and Economic Statistics*, 15:68-73.

Christiano, Lawrence J. and Martin Eichenbaum, 1990, Unit roots in real GNP: Do we know and do we care?, in *Unit Roots, Investment Measures, and Other Essays*, ed. by Allen H. Meltzer, Carnegie-Rochester Conference Series on Public Policy, 32:7-61.

Clark, Peter K., 1987, The cyclical component of U.S. economic activity, *Quarterly Journal of Economics*, 102:797-814.

Cochrane, John H., 1988, How big is the random walk in GNP?, *Journal of Political Economy*, 96:893-920.

Cochrane, John H., 1991, A critique of the application of unit root tests, *Journal of Economic Dynamics and Control*, 15:275-84.

Davidson, Russell, and James G. MacKinnon, 1993, *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

DeLong, J. B. and L. H. Summers, 1986, Are business cycles symmetrical?, in: R. J. Gordon, ed., *The American Business Cycle: Continuity and Change*, University of Chicago Press, Chicago.

DeLong, J. B. and L. H. Summers, 1988, How does macroeconomic policy affect output?, *Brookings Papers on Economic Activity*, 0:433-480.

Dickey, David A., 1976, Estimation and hypothesis testing in nonstationary time series, Unpublished doctoral dissertation, Iowa State University.

Dickey, David A., William R. Bell and Robert B. Miller, 1986, Unit roots in time series models: Tests and implications, *American Statistician*, 40:12-26.

Dickey, David A. and Wayne A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74:427-431.

Diebold, F. X. and G. D. Rudebusch, 1990, A nonparametric investigation of duration dependence in the American business cycle, *Journal of Political Economy*, 98: 596-616.

Diebold, F. X. and G. D. Rudebusch, 1996, Measuring business cycles: A modern perspective, *The Review of Economics and Statistics*, 78:67-77.

Diebold, F. X., G. D. Rudebusch, and D. E. Sichel, 1993, Further evidence on business cycle duration dependence, in: J. H. Stock and M. W. Watson, eds., *Business cycles, Indicators, and Forecasting*, University of Chicago Press, Chicago, 255-280.

Diebold, Francis X., and Abdelhak S. Senhadji, 1996, The uncertain unit root in real GNP: Comment, *American Economic Review*, 86:1291-98.

Durland, J. M. and T. H. McCurdy, 1994, Duration-dependent transitions in a Markov model of U. S. GNP growth, *Journal of Business and Economic Statistics*, 12: 279-288.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock, 1996, Efficient tests for an autoregressive unit root, *Econometrica*, 64:813-36.

Engel, Charles, 1997, Long-run PPP may not hold after all, working paper, University of Washington.

Engle, Robert F., and C. W. J. Granger, 1987, Cointegration and error correction: Representation, estimation, and testing, *Econometrica*, 55:251-76.

Falk, Barry, 1986, Further evidence on the asymmetric behavior of economic time series over the business cycle, *Journal of Political Economy*, 94:1096-1109.

Frances, Philip H., and Niels Haldrup, 1994, The effects of additive outliers on tests for unit roots and cointegration, *Journal of Business and Economic Statistics*, 12:471-78.

Friedman, Milton, 1964, Monetary studies of the national bureau, the national bureau enters its 45th year, 44th annual report, 7-25; Reprinted in M. Friedman, 1969, *The Optimum Quantity of Money and Other Essays*, Aldine: Chicago, 261-284.

Friedman, Milton, 1993, The 'plucking model' of business fluctuations revisited, *Economic Inquiry*, 31:171-177.

Friedman, Milton and Anna J. Schwartz, 1963, *A Monetary History of the United States, 1867-1960*, NBER, New York.

Fuller, W.A., 1976, *Introduction to Statistical Time Series*, Wiley, New York.

Gallman, Robert E. 1966, Gross national product in the United States 1834-1909, in *Studies in Income and Wealth*, vol. 30, Columbia University Press (for NBER), New York.

Gordon, K. and A. F. M. Smith, 1988, Modeling and monitoring discontinuous changes in time series, in: J. C. Spall, ed., *Bayesian Analysis of Time Series and Dynamic Linear Models*, Marcel Dekker, New York, 359-392.

Gordon, Robert J., 1993, The jobless recovery: Does it signal a new era of productivity-led growth?, *Brookings Papers on Economic Activity*, 0:271-306.

Hall, Alastair, 1994, Testing for a unit root in time series with pretest data-based model selection, *Journal of Business and Economic Statistics*, 12:461-70.

Hall, Alastair R., Glenn D. Rudebusch, and David W. Wilcox, 1994, Judging instrument relevance in instrumental variables estimation, *International Economic Review*, 37:283-98.

Hamilton, James D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57:357-384.

Hamilton, James D., 1994, *Time Series Analysis*, Princeton University Press, Princeton.

Hansen, Bruce E., 1996, Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica*, 64:413-430.

Harrison, P. J. and C. F. Stevens, 1976, Bayesian forecasting, *Journal of the Royal Statistical Society*, Ser. B, 38:205-247.

Jones, Charles I., 1995, Time series tests of endogenous growth models, *Quarterly Journal of Economics*, 110:495-525.

Journal of Applied Econometrics, 1992, 7(S).

Journal of Econometrics, 1996, 74(1.)

Kendrick, John W., 1961, *Productivity Trends in the United States*, Princeton University Press, Princeton

Keynes, J. M., 1936, *The General Theory of Employment, Interest and Money*, Macmillan, London.

Kilian, Lutz and Lee E. Ohanian, 1996, "Is there a trend break in U.S. GNP? A macroeconomic perspective, unpublished manuscript.

Kim, Chang-Jin, 1993, Unobserved-component time series models with Markov-switching heteroskedasticity: Changes in regime and the link between inflation rates and inflation uncertainty, *Journal of Business and Economic Statistics*, 11:341-349.

Kim, Chang-Jin, 1994, Dynamic linear models with Markov switching, *Journal of Econometrics*, 60:1-22.

Kim, Chang-Jin and C. R. Nelson, 1998a, Business cycle turning points, a new coincident index, and tests for duration dependence based on a dynamic factor model with Markov switching, *The Review of Economics and Statistics*, 80:188-201.

Kim, Chang-Jin and C. R. Nelson, 1998b, Friedman's plucking model of business fluctuations: Tests and estimates of permanent and transitory components, working paper, University of Washington discussion paper.

Kim, M.-J. and J.-S. Yoo, 1995, New index of coincident indicators: A multivariate Markov switching factor model approach, *Journal of Monetary Economics*, 36: 607-630.

Kuznets, Simon S., 1941, *National Income and Its Composition, 1919-1938*, NBER, New York.

Kuznets, Simon S., 1946, *National Product Since 1869*, NBER, New York.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," *Journal of Econometrics*, 54:159-178.

Lam, Pok-Sang., 1990, The Hamilton model with a general autoregressive component: Estimation and comparison with other models of economic time series, *Journal of Monetary Economics*, 26: 409-432.

Leybourne S. J., and McCabe, B. P. M., 1994, A consistent test for a unit root, *Journal of Business and Economic Statistics*, 12:157-66.

Maddison, Angus, 1995, *Monitoring the World Economy 1820-1992*, OECD, Paris.

Mitchell, W. A., 1927, *Business Cycles: The Problem and its Setting*, NBER, New York.

Murray, Christian J., 1997, Trend stationarity in annual U.S. RGDP? A reexamination of recent evidence, working paper, University of Washington.

Murray, Christian J. and Charles R. Nelson, 1998, The uncertain trend in U.S. GDP, working paper, University of Washington.

Neftçi, Salih. N., 1984, Are economic time series asymmetric over the business cycle?, *Journal of Political Economy*, 92:307-328.

Nelson, Charles R. and Heejoon Kang, 1981, Spurious periodicity in inappropriately detrended time series, *Econometrica*, 49:741-51.

Nelson, Charles R. and Charles I. Plosser, 1982, Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of Monetary Economics*, 10:139-162.

Ng, Serena, and Pierre Perron, 1995, Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag, *Journal of the American Statistical Association*, 90:268-81.

Perron, Pierre, 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57:1361-1401.

Perron, Pierre, 1997, Further evidence on breaking trend functions in macroeconomic variables, *Journal of Econometrics*, 80:355-385.

Perron, Pierre and Timothy Vogelsang, 1992, Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business and Economic Statistics*, 10:301-320.

Phillips, P. C. B. and Pierre Perron, 1988, Testing for a unit root in time series regression, *Biometrika*, 75:335-46.

Potter, Simon. M., 1995, A nonlinear approach to U. S. GNP, *Journal of Applied Econometrics*, 10:109-125.

Rappoport, P. and L. Reichlin, 1989, Segmented trends and nonstationary time series, *The Economic Journal*, 99:168-177.

Romer, Christina D., 1989, The prewar business cycle reconsidered, new estimates of gross national product 1869-1908, *Journal of Political Economy*, 97:1-37.

Rudebusch, Glenn D., 1992, Trends and random walks in macroeconomic time series: A re-examination, *International Economic Review*, 33:661-80.

Rudebusch, Glenn D., 1993, The uncertain unit root in real GNP, *American Economic Review*, 83:264-72.

Schwartz, G., 1978, Estimating the dimension of a model, *The Annals of Statistics*, 6: 461-64.

Schwert, G. William, 1987, Effects of model specification on tests for unit roots in macroeconomic data, *Journal of Monetary Economics*, 20:73-103.

Schwert, G. William, 1989a, Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, 7:147-60.

Schwert, G. William, 1989b, Why does stock market volatility change over time?, *Journal of Finance*, 44:1115-53.

Sichel, Daniel E., 1989, Are business cycles asymmetric? A correction, *Journal of Political Economy*, 97:1055-1060.

Sichel, Daniel E., 1991, Business cycle duration dependence: A parametric approach, *The Review of Economics and Statistics*, 73:254-260.

Sichel, Daniel E., 1993, Business cycle asymmetry: A deeper look, *Economic Inquiry*, 31:224-236.

Sichel, Daniel E., 1994, Inventories and the three phases of the business cycle, *Journal of Business and Economic Statistics*, 12:269-277.

Stock, James H., 1991, Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series, *Journal of Monetary Economics*, 28:435-459.

Stock, James H. and Mark W. Watson, 1989, New indexes of coincident and leading indicators, in *NBER Macroeconomic Annual*, eds. O. J. Blanchard and S. Fischer, MIT Press, Cambridge, 351-393.

Stock, James H. and Mark W. Watson, 1991, A probability model of the coincident economic indicators, in: K. Lahiri and G. H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge University Press, New York, 63-85.

Stock, James H. and Mark W. Watson, 1993, A procedure for predicting recessions with Leading indicators: Econometric issues and recent experiences, in: James H. Stock and Mark W. Watson, eds., *Business Cycles, Indicators, and Forecasting* University of Chicago Press, Chicago, 95-156.

Survey of Current Business, 1996, volume 76, (January/February), U.S. Department of Commerce.

Vogelsang, Timothy J., 1993, Essays on testing for nonstationarities and structural change in time series models, Unpublished Ph.D. Dissertation, Princeton University, Princeton, NJ.

Vogelsang, Timothy, J., 1997, Wald-type tests for detecting breaks in the trend function of a dynamic time series, *Econometric Theory*, 13:818-849.

Vogelsang, Timothy, J. and Pierre Perron, 1994, Additional tests for a unit root allowing or a break in the trend function at an unknown time, working paper, Cornell University.

Wynne, Mark A. and Nathan S. Balke, 1992, Are deep recessions followed by strong recoveries?, *Economics Letters*, 39:183-189.

Zivot, Eric and Donald W. K. Andrews, 1992, Further evidence on the great crash, the oil-price shock, and the unit root hypothesis, *Journal of Business and Economic Statistics*, 10:251-270.

APPENDIX A: IRREGULAR UNOBSERVED COMPONENTS MODEL FOR GDP

Here we present the Markov switching model discussed in Section 1.2. The model decomposes the natural log of GDP into an unobserved trend, cycle, and irregular component. It is a generalization of Clark's (1987) model. Letting y_t denote the log of GDP, we have:

$$y_t = \eta_t + c_t + S_t i_t$$

$$\eta_t = g_{t-1} + \eta_{t-1} + v_t$$

$$g_t = g_{t-1} + w_t$$

$$\psi(L)c_t = \varepsilon_t$$

$$i_t = \phi_1 i_{t-1} + \phi_2 i_{t-2} + u_t$$

$$\begin{bmatrix} v_t \\ w_t \\ \varepsilon_t \\ u_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & 0 & 0 & 0 \\ 0 & \sigma_w^2 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_u^2 \end{bmatrix} \right)$$

$$S_t = \{0,1\};\ \Pr[S_t = 0 \mid S_{t-1} = 0] = q;\ \Pr[S_t = 1 \mid S_{t-1} = 1] = p.$$

This model reduces to Clark's model when S_t=0. To anticipate our results, when we allow for the irregular component, i_t, the cyclical component becomes irrelevant. Thus we specify $\psi(L)$=0. Writing the model in state space for, we have:

$$y_t = H_{S_t} \xi_t$$

$$\xi_t = F_{S_t}\xi_{t-1} + \upsilon_t$$

$$E(\upsilon_t \upsilon_t^{'}) = Q_{S_t}$$

where

$$\xi_t = [\eta_t \quad g_t \quad i_t \quad i_{t-1}]'$$

$$\upsilon_t = [v_t \quad w_t \quad u_t \quad 0]'$$

$$H_{S_t} = [1 \quad 0 \quad S_t \quad 0]$$

$$F_{S_t} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & \phi_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$Q_{S_t} = \begin{bmatrix} \sigma_v^2 & 0 & 0 & 0 \\ 0 & \sigma_w^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Estimation of the parameters, as well as the unobserved components, is based on Kim's (1994) approximate maximum likelihood algorithm. Briefly, state space models with Markov switching render computation of the exact likelihood function via the Kalman filter intractable. Kim modifies the Kalman filtering equations to render construction of the Gaussian likelihood feasible. Since the state vector, ξ_t , is nonstationary, we discard the first nine sample points so that the initial guesses for the state vector and mean-squared error are allowed to die out. The plot in Figure 1.3 is from 1889-1997. Parameter estimates are as follows:

Parameter	q	p	ϕ_1	ϕ_2	σ_v	σ_u	σ_w
Estimate	0.9811 (0.0230)	0.9776 (0.0191)	1.2467 (0.1277)	-0.3886 (0.0796)	0.0286 (0.0000)	0.0581 (0.0002)	0.0000 (0.0000)

Standard errors are in parenthesis.

APPENDIX B: STATE SPACE MODELING AND KALMAN FILTERING FOR THE 2-STATE MODEL

In this section, we discuss representation and estimation of Model 4, based on Kim's (1994) method of approximate maximum likelihood estimation. We employ the following state space representation for equations (2.3.10)'-(2.3.14)', and (2.3.15)-(2.3.17):

Measurement equation: $\Delta y_t = H\xi_t$

Transition equation: $\xi_t = \mu_{S_t} + F\xi_{t-1} + V_t$

$$E(V_t V_t') = Q,$$

where

$$H = \begin{bmatrix} \gamma_1 & 0 & \lambda_1 & -\lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \gamma_2 & 0 & \lambda_2 & -\lambda_2 & 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma_3 & 0 & \lambda_3 & -\lambda_3 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\xi_t = \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ x_t \\ x_{t-1} \\ z_{1t} \\ z_{1,t-1} \\ z_{2t} \\ z_{2,t-1} \\ z_{3t} \\ z_{3,t-1} \end{bmatrix}, \quad \mu_{S_t} = \begin{bmatrix} \beta_{S_t} \\ 0 \\ \tau S_t \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad V_t = \begin{bmatrix} v_t \\ 0 \\ u_t \\ 0 \\ e_{1t} \\ 0 \\ e_{2t} \\ 0 \\ e_{3t} \\ 0 \end{bmatrix},$$

$$F = \begin{bmatrix} \phi_1 & \phi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_1^{\bullet} & \phi_2^{\bullet} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \psi_{11} & \psi_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \psi_{21} & \psi_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{31} & \psi_{32} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then, conditional on $S_t = j$ and $S_{t-1} = i$, the Kalman filter equations can be written as:

$$\xi_{t|t-1}^{(i,j)} = \mu_{S_j} + F\xi_{t-1|t-1}^{i},$$

$$P_{t|t-1}^{(i,j)} = FP_{t-1|t-1}^{i}F' + Q,$$

$$\eta_{t|t-1}^{(i,j)} = \Delta y_t - H\xi_{t|t-1}^{(i,j)},$$

$$f_{t|t-1}^{(i,j)} = HP_{t|t-1}^{(i,j)}H',$$

$$\xi_{t|t}^{(i,j)} = \xi_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)}H'[f_{t|t-1}^{(i,j)}]^{-1}\eta_{t|t-1}^{(i,j)},$$

$$P_{t|t}^{(i,j)} = (I - P_{t|t-1}^{(i,j)} H' [f_{t|t-1}^{(i,j)}]^{-1}) H P_{t|t-1}^{(i,j)},$$

where $\xi_{t|t-1}^{(i,j)}$ is an inference on ξ_t based on information up to time t-1, conditional on $S_t = j$ and $S_{t-1} = i$; $\xi_{t|t}^{(i,j)}$ is an inference on ξ_t based on information up to time t, conditional on $S_t = j$ and $S_{t-1} = i$; $P_{t|t-1}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ are the MSE matrices of $\xi_{t|t-1}^{(i,j)}$ and $\xi_{t|t}^{(i,j)}$ respectively; $\eta_{t|t-1}^{(i,j)}$ is the conditional forecast error of Δy_t based on information up to time t-1; $f_{t|t-1}^{(i,j)}$ is the conditional variance of $\eta_{t|t-1}^{(i,j)}$.

As noted by Harrison and Stevens (1976) and Gordon and Smith (1988) each iteration produces a two-fold increase in the number of cases to consider. To render the Kalman filter operational, we need to collapse the 2^2 posteriors ($\xi_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$) into 2 at each iteration. Collapsing requires the following approximations suggested by Harrison and Stevens (1976):

$$\xi_{t|t}^{j} = \frac{\sum_{i=0}^{1} \Pr[S_{t-1} = i, S_t = j \mid \psi_t] \xi_{t|t}^{(i,j)}}{\Pr[S_t = j \mid \psi_t]},$$

and

$$P_{t|t}^{j} = \frac{\sum_{i=0}^{1} \Pr[S_{t-1} = i, S_t = j \mid \psi_t] \left\{ P_{t|t}^{(i,j)} + (\xi_{t|t}^{j} - \xi_{t|t}^{(i,j)})(\xi_{t|t}^{j} - \xi_{t|t}^{(i,j)})' \right\}}{\Pr[S_t = j \mid \psi_t]},$$

where ψ_t refers to information available at time t.

In order to obtain the probability terms necessary for collapsing, we perform the following procedure due to Hamilton (1989):

Step 1:

At the beginning of the t^{th} iteration, given $\Pr[S_{t-1} = i \mid \psi_{t-1}]$ for $i = 0$ or 1, we can calculate

$$\Pr[S_t = j, S_{t-1} = i \mid \psi_{t-1}] = \Pr[S_t = j \mid S_{t-1} = i] \Pr[S_{t-1} = i \mid \psi_t],$$

$$(\mathrm{i}, \mathrm{j} = 0, 1).$$

Step 2:

Consider the joint density of Δy_t, S_t, and S_{t-1} :

$$f(\Delta y_t, S_t = j, S_{t-1} = i \mid \psi_{t-1}) = f(\Delta y_t \mid S_t = j, S_{t-1} = i, \psi_{t-1}) \Pr[S_t = j, S_{t-1} = i \mid \psi_{t-1}]$$

from which the marginal density of Δy_t is obtained by:

$$f(\Delta y_t \mid \psi_{t-1}) = \sum_{i=0}^{1} \sum_{j=0}^{1} f(\Delta y_t, S_t = j, S_{t-1} = i \mid \psi_{t-1})$$

$$= \sum_{i=0}^{1} \sum_{j=0}^{1} f(\Delta y_t \mid S_t = j, S_{t-1} = i, \psi_{t-1}) \Pr[S_t = j, S_{t-1} = i \mid \psi_{t-1}]$$

where the conditional density $f(\Delta y_t \mid S_t = j, S_{t-1} = i, \psi_{t-1})$ is obtained via the prediction-error decomposition:

$$f(\Delta y_t \mid S_t = j, S_{t-1} = i, \psi_{t-1}) = (2\pi)^{-\frac{T}{2}} \left| f_{t|t-1}^{(i,j)} \right|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \eta_{t|t-1}^{(i,j)\prime} f_{t|t-1}^{(i,j)^{-1}} \eta_{t|t-1}^{(i,j)} \right\}.$$

Step 3:

Once Δy_t is observed at the end of time t, we update the probability terms:

$$\Pr[S_t = j, S_{t-1} = i \mid \psi_t]$$

$$= \Pr[S_t = j, S_{t-1} = i \mid \psi_{t-1}, \Delta y_t]$$

$$= \frac{f(S_t = j, S_{t-1} = i, \Delta y_t \mid \psi_{t-1})}{f(\Delta y_t \mid \psi_{t-1})}$$

$$= \frac{f(\Delta y_t \mid S_t = j, S_{t-1} = i, \psi_{t-1}) \Pr[S_t = j, S_{t-1} = i \mid \psi_{t-1}]}{f(\Delta y_t \mid \psi_{t-1})}$$

with

$$\Pr[S_t = j \mid \psi_t] \sum_{i=0}^{1} \Pr[S_t = j, S_{t-1} = i \mid \psi_t].$$

To initialize the above filter, we use the steady-state probabilities given by[9]

$$\Pr[S_0 = 0] = \frac{1-p}{2-p-q} \quad \text{and} \quad \Pr[S_0 = 1] = \frac{1-q}{2-p-q}.$$

As a by product of the above filter, we obtain the log likelihood function:

[9] See Hamilton (1994, p.684)

$$\ln L = \sum_{t=1}^{T} ln(f(\Delta y_t \mid \psi_{t-1})$$

which can be maximized with respect to the parameters of the model.

To calculate the level of the common permanent component, proceed as follows. Since the data are in deviations from their means, δ and $D = [D_1\ D_2\ D_3]'$ are concentrated out of the likelihood function. As in Stock and Watson (1991), we can use the steady state Kalman gain retrieve these terms in the following manner:

$$\delta = E'(I_r - (I_r - K^* H)F)^{-1} K^* \Delta\bar{y},$$

$$\tilde{D} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \Delta\bar{y} - \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} \delta,$$

where K^* is the steady state Kalman gain[10], $E_1' = [1\,0\,0 \ldots 0]'$, and r is the dimension of the state vector. Once δ is retrieved, given $\Delta\tilde{c}_T = [\Delta c_1\ \Delta c_2 \ldots \Delta c_T]'$, and arbitrary initial value C_0, we obtain $C_t = \delta + \Delta c_t + C_{t-1}$, $t = 1,2,\ldots,T$.

[10] At each iteration of the Kalman filter, there are 4 Kalman gain matrices. The steady state gain is calculated as the weighted mean of the 4 gain matrices in the final iteration.

APPENDIX C: KALMAN FILTERING FOR THE 4-STATE MODEL

The state space representation of Model 5 differs from that of Model 4 in that μ_{S_t} is now $\mu_{S_{1t},S_{2t}} = [\beta_{S_{1t}} \quad 0 \quad \tau S_{2t} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]'$. Conditional on $S_{1t} = j, S_{2t} = j', S_{1,t-1} = i$, and $S_{2,t-1} = i'$, the Kalman filtering equations can be written as:

$$\xi_{t|t-1}^{(i,i',j,j')} = \mu_{S_{1j},S_{2j'}} + F\xi_{t-1|t-1}^{(i,i')},$$

$$P_{t|t-1}^{(i,i',j,j')} = FP_{t-1|t-1}^{(i,i')}F' + Q,$$

$$\eta_{t|t-1}^{(i,i',j,j')} = \Delta y_t - H\xi_{t|t-1}^{(i,i',j,j')},$$

$$f_{t|t-1}^{(i,i',j,j')} = HP_{t|t-1}^{(i,i',j,j')}H',$$

$$\xi_{t|t}^{(i,i',j,j')} = \xi_{t|t-1}^{(i,i',j,j')} + P_{t|t-1}^{(i,i',j,j')}H'[f_{t|t-1}^{(i,i',j,j')}]^{-1}\eta_{t|t-1}^{(i,i',j,j')},$$

$$P_{t|t}^{(i,i',j,j')} = (I - P_{t|t-1}^{(i,i',j,j')}H'[f_{t|t-1}^{(i,i',j,j')}]^{-1})HP_{t|t-1}^{(i,i',j,j')}.$$

At each iteration, we collapse the 4^2 terms into 4 as follows:

$$\xi_{t|t}^{(j,j')} = \frac{\sum_{i=0}^{1}\sum_{i'=0}^{1}\Pr[S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i' \mid \psi_t]\xi_{t|t}^{(i,i',j,j')}}{\Pr[S_{1t} = j, S_{2t} = j' \mid \psi_t]}$$

and

$$P_{t|t}^{(j,j')} = \frac{\left\{\begin{matrix}\sum_{i=0}^{1}\sum_{i'=0}^{1}\Pr[S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i' \mid \psi_t] \\ \times [P_{t|t}^{(i,i',j,j')} + (\xi_{t|t}^{(j,j')} - \xi_{t|t}^{(i,i',j,j')})(\xi_{t|t}^{(j,j')} - \xi_{t|t}^{(i,i',j,j')})']\end{matrix}\right\}}{\Pr[S_{1t} = j, S_{2t} = j' \mid \psi_t]}.$$

To obtain the probability terms we employ Hamilton's (1989) filter again:

Step 1:

At the beginning of the t^{th} iteration, given $\Pr[S_{1,t-1} = i \mid \psi_{t-1}]$ and $\Pr[S_{2,t-1} = i' \mid \psi_{t-1}]$ for $i, i' = 0$ or 1, we can calculate

$$\Pr[S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1}]$$

$$=\Pr[S_{1t}=j\mid S_{1,t-1}=i]\Pr[S_{1,t-1}=i\mid\psi_{t-1}]\Pr[S_{2t}=j'\mid S_{2,t-1}=i']\Pr[S_{2,t-1}=i'\mid\psi_{t-1}],$$

$$(i,i',j,j'=0,1).$$

<u>Step 2:</u>

Consider the joint density of $\Delta y_t, S_{1t}, S_{1,t-1}, S_{2t}$, and $S_{2,t-1}$:

$$f(\Delta y_t,S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1})$$
$$=f(\Delta y_t\mid S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i',\psi_{t-1})$$
$$\times\ \Pr[S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1}]$$

from which the marginal density of Δy_t is obtained by:

$$f(\Delta y_t\mid\psi_{t-1})=\sum_{i=0}^{1}\sum_{i'=0}^{1}\sum_{j=0}^{1}\sum_{j'=0}^{1}f(\Delta y_t,S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1})$$

$$=f(\Delta y_t\mid S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i',\psi_{t-1})$$
$$\times\ \Pr[S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1}]$$

where the conditional density of Δy_t is obtained via the prediction-error decomposition:

$$f(\Delta y_t\mid S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i',\psi_{t-1}),$$

$$=(2\pi)^{-\frac{T}{2}}\left|f_{t|t-1}^{(i,i',j,j')}\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\eta_{t|t-1}^{(i,i',j,j')\prime}f_{t|t-1}^{(i,i',j,j')^{-1}}\eta_{t|t-1}^{(i,i',j,j')}\right\}$$

<u>Step 3:</u>

Once Δy_t is observed at the end of time t, we update the probability terms:

$$\Pr[S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_t]$$

$$=\Pr[S_{1t}=j\mid S_{1,t-1}=i]\Pr[S_{1,t-1}=i\mid\psi_t]\Pr[S_{2t}=j'\mid S_{2,t-1}=i']\Pr[S_{2,t-1}=i'\mid\psi_{t-1},\Delta y_t]$$

$$=\frac{f(\Delta y_t,S_{1t}=j,S_{2t}=j',S_{1,t-1}=i,S_{2,t-1}=i'\mid\psi_{t-1})}{f(\Delta y_t\mid\psi_{t-1})}$$

$$= \begin{pmatrix} f(\Delta y_t \mid S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i', \psi_{t-1}) \\ \times \ \Pr[S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i' \mid \psi_{t-1}] \end{pmatrix} \div f(\Delta y_t \mid \psi_{t-1})$$

with

$$\Pr[S_{1t} = j, S_{1,t-1} = i \mid \psi_t] = \sum_{i'=0}^{1} \sum_{j'=0}^{1} \Pr[S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i' \mid \psi_t]$$

and

$$\Pr[S_{2t} = j', S_{2,t-1} = i' \mid \psi_t] = \sum_{i=0}^{1} \sum_{j=0}^{1} \Pr[S_{1t} = j, S_{2t} = j', S_{1,t-1} = i, S_{2,t-1} = i' \mid \psi_t].$$

Finally,

$$\Pr[S_{1t} = j \mid \psi_t] = \sum_{i=0}^{1} \Pr[S_{1t} = j, S_{1,t-1} = i \mid \psi_t],$$

and

$$\Pr[S_{2t} = j' \mid \psi_t] = \sum_{i'=0}^{1} \Pr[S_{2t} = j', S_{2,t-1} = i' \mid \psi_t].$$

VITA

Christian Joseph Murray

University of Washington

1998

I was born in December of 1970 in Seattle, Washington. I received a Bachelor of Arts in Economics from Pacific Lutheran University in May of 1993. I began my graduate studies at the University of Washington in September of 1994. I received a Master of Arts in Economics in the summer of 1996. I received my Doctoral degree in Economics in June of 1998.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989