Trend and Cycles: A New Approach and Explanations of Some Old Puzzles*

Pierre Perron[†] **Tatsuma Wada**[‡]

Boston University **Boston University**

This Version: January 2, 2005

Abstract

Recent work on trend-cycle decompositions for US real GDP yield the following puzzling features: method based on Unobserved Components models, the Beveridge-Nelson decomposition, the Hodrick-Prescott filter and others yield very different cycles which bears little resemblance to the NBER chronology, ascribes much movements to the trend leaving little to the cycle, and some imply a negative correlation between the noise to the cycle and the trend. We argue that these features are artifacts created by the neglect of the presence of a change in the slope of the trend function in real GDP in 1973. Once this is properly accounted for, the results show all methods to yield the same cycle with a trend that is non-stochastic except for a few periods around 1973. This cycle is more important in magnitude than previously reported, it accords very well with the NBER chronology and imply no correlation between the trend and cycle, since the former is non-stochastic. We propose a new approach to univariate trend-cycle decompositions using a generalized Unobserved Components models with errors having a mixture of Normals distribution for both the slope of the trend function and the cyclical component. It can account endogenously for infrequent changes such as level shifts and change in slope, as well as different variances for expansions and recessions. It yields a decomposition that accords very well with common notions of the business cycle.

JEL Classification Number: C22, E32.

Keywords: Trend-Cycle Decomposition, Structural Change, Non Gaussian Filtering, Unobserved Components Model, Beveridge-Nelson Decomposition.

*We are grateful to Zhongjun Qu for helpful discussions.
[†]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).
[‡]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (twada@bu.edu).

1 Introduction

Interest in the business cycle has a long standing history in both theoretical investigations and empirical applications. The important contribution of Burns and Mitchell (1946) paved the way for methods to measure it. The literature has, however, departed from their methods due to its complexity and the need for subjective evaluations. Instead, much of the work has concentrated on easily applicable mechanical and non-subjective methods. In the last few decades, many alternative procedures have been suggested. The need for a quantitative measure of the business cycle arises in great part because most macroeconomic models delivers implications that pertains to the non-trending component of series. In order to confront these models with the data, there is, accordingly, a need to separate the trend and the cycle. This issue of trend-cycle decomposition is the object of our paper.

We shall concentrate on the trend-cycle decomposition of (log) post-war quarterly US real GDP seasonally adjusted. Popular methods to extract the cyclical component include the following, among others: the Beveridge-Nelson (1981) (BN) decomposition based on unconstrained ARIMA model (Campbell and Watson, 1987, Watson, 1986, Cochrane, 1986), Unobserved Components (UC) (Clark, 1987), the Hodrick-Prescott (1997) (HP) filter, and the Band-Pass (BP) filter (Baxter and King, 1999). Note that the latter is not per se a trend-cycle decomposition since it also eliminates high-frequency components. Nevertheless, for the real GDP series analyzed, high frequency movements are not important and it effectively acts as a trend removal procedure. For reviews and applications, see Stock and Watson (1987, 1999).

A major problem faced by practitioners is that these methods usually lead to a different trend cycle decomposition and the differences are often substantial and they lead to quite different "stylized facts" about the business cycle to be used when confronting models with the data (see, e.g., Canova, 1998, 1999). It is therefore important to carefully assess the suitability of each method. In this paper, we shall concentrate on the BN and UC decompositions with some remarks about the HP and BP filters.

It is well known that the UC and BN decompositions yield very different cycles, in particular the latter ascribes most movements to the trend and leaves little to the cycle. Such differences may, at first sight, not be surprising since the UC decomposition assumes no correlation between the shock to the trend and the cycle, while the BN decomposition assumes a perfect correlation (provided the trend is stochastic). In a recent paper, Morley et al. (2003) show that by specifying a simple AR(2) process for the cycle, it is possible to identify an unobserved components model in which this correlation is a free parameter to be estimated. When doing so, the data suggest a high (negative) correlation and the decomposition is virtually identical to that provided by BN.

Hence, we are left with what may be perceived by some as a set of puzzling features: 1) the fact that these methods yield drastically different answers; 2) the small scale and noisy structure of the cycle delivered by the BN decomposition; 3) the fact that most methods yield a cycle that bears little resemblance to the NBER chronology; 4) the negative correlation between the noise to the cycle and the trend.

We shall argue that these puzzling features are artifacts created by the neglect of the presence of a change in the slope of the trend function in real GDP. Once this is properly accounted for the results show that 1) all methods yields the same cycle and the trend is non-stochastic except for a few periods around the date of the change in the slope; 2) this cycle is important in magnitude, more so than previously reported; 3) it accords very well with the NBER chronology; 4) there is no correlation between the trend and cycle, since the former is non-stochastic. All the puzzling features disappears and we are left with a cyclical component that agrees much better with common notions of the business cycle.

The outline of the paper is as follows. Section 2 presents preliminary results about the trend-cycle obtained using standard unobserved component models (with and without correlation in the noise of the trend and cycle) and their relation to the Beveridge-Nelson decomposition. We discuss the important differences across various specifications. Section 3 presents similar decompositions for which the only modification is to allow for the possibility of a one time change in the slope of the trend function in 1973:1. The results are very different from those models without the possibility of a change in slope, yet they all agree across different specifications. Section 4 shows via simulations that if our decomposition is right, previous results can be explained, in particular the importance of variations in the trend component and the negative correlation between the noise to the trend and the noise to the cycle. Section 5 presents an alternative framework for trend-cycle decompositions based on mixtures of Normal distributions for the noise components. It is able to capture infrequent changes to the slope and also allows different variances in the cycle for periods of recessions and expansions. The results show a trend-cycle decomposition that agrees well with common notions of business cycles and the NBER chronology. Section 6 presents further comparisons with the Hodrick-Prescott and Band Pass filters. Section 7 offers brief conclusions. An appendix contains all technical material related to estimations and simulations performed. We organized the paper such that the main text contains only a description of the main issues and results without the important technical details contained in the appendix. Note finally that since we shall make frequent comparisons to the results of Morley et al. (2003), we use exactly the same data set that they used, namely the (log) quarterly US real GDP series seasonally adjusted for the period 1947:1-1998:2. Hence, any differences reported will be the result of different methods, not different data.

2 Preliminaries

Consider the basic unobserved components model that describes log real GDP y_t as the sum of a trend τ_t and a cyclical component c_t:

$$\begin{aligned} y_t &= \tau_t + c_t \\ \tau_t &= \mu + \tau_{t-1} + \eta_t \\ A(L)c_t &= B(L)\epsilon_t \end{aligned} \tag{1}$$

where $A(L)$ and $B(L)$ are polynomials in L of order p and q, respectively, with all roots outside the unit circle, and

$$\begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix} \sim i.i.d.\ N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\eta^2 & \sigma_{\eta\epsilon} \\ \sigma_{\eta\epsilon} & \sigma_\epsilon^2 \end{bmatrix}\right).$$

Hence, the trend is a random walk with drift and the cycle is an $ARMA(p,q)$ process. It is well known that, under this level of generality, the model is not identified (see, e.g., Watson, 1986). A sufficient condition for identification is to specify a value for the covariance $\sigma_{\eta\epsilon}$. A popular choice is to set $\sigma_{\eta\epsilon} = 0$, that is to specify that the shocks to the trend function are uncorrelated with the shocks to the cyclical component. This implies that the reduced form of the system is a constrained $ARMA(p,q^*)$ with $q^* = \max(p, q+1)$. In particular, the constraints are such that the only class of permissible models are those for which the spectral density function of Δy_t, the growth rates of real GDP, takes a minimal value at frequency zero. This rules out, in particular, an AR(1) specification for Δy_t. This specification will be denoted $UC0$.

An alternative specification is to assume that the shocks are perfectly (negatively) correlated. In this case the reduced form is an unconstrained $ARMA(p,q^*)$ process of the form

$$A(L)\Delta y_t = \mu + B^*(L)u_t$$

where $u_t \sim i.i.d.\ N(0, \sigma_u^2)$ with the value σ_u^2 depending on the parameters of the model. The trend function can then equivalently be obtained via the Beveridge-Nelson (1981) decomposition

$$BN_t = \mu + BN_{t-1} + \varphi(1)u_t$$

with $\varphi(1) = B^*(1)/A(1)$. This decomposition will be denoted with the acronym BN.

To see the trend and cycle decompositions implied by each specification, we use the same data set as in Morley et al. (2003), namely the logarithm of U.S. real GDP 1947:1-1998:2 seasonally adjusted. A specification for the cyclical component that was found to be adequate is a simple AR(2), and accordingly an ARIMA(2,1,2) for the BN decomposition is used. Figures 1 reproduces the results of Morley et al. (2003) for the estimated trend and cycle for each method (details about the estimation method can be found in Appendix A). As can be seen the decompositions are very different. The BN decomposition ascribes most movements in the real GDP series to the trend function leaving a cyclical component that is very small, noisy and which bears no resemblance to the NBER chronology, whose periods of recessions are indicated with a shaded area. On the other hand, the $UC0$ decomposition leaves more importance to the cyclical component whose peaks and throughs corresponds somewhat more closely to the NBER chronology.

The great differences in the implied trend and cycle suggests that either or both of the crucial identifying assumptions are at odds with the data, i.e., the correlation between the shocks to the trend and the cycle may not be 0 or 1. With cleaver insight, Morley et al. (2003) recognized that it is possible to identify model (1) with $\sigma_{\eta\epsilon}$ unconstrained, provided $p \geq q + 2$. With an $AR(2)$ cyclical component ($p = 2, q = 0$), we have a just identified system. This is important since an $AR(2)$ specification for the cycle is, at least with US data, a reasonable approximation which can be tested ex-post. Following Morley et al. (2003), this decomposition is labeled $UCUR$. The resulting trend-cycle decomposition is found to be indistinguishable from the BN decomposition (hence, the graphs are not repeated). This suggests the following conclusions: 1) the data do not support the hypothesis that the correlation between the shocks to the trend and the cycle is 0; 2) the constraints imposed by the $UCUR$ model are compatible with the estimated unrestricted $ARIMA(2,1,2)$ model.

Remark 1 *It is important to note that the results discussed so far remain basically unchanged if the trend function is specified as follows*

$$\begin{aligned}\tau_t &= \beta_t + \tau_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + \omega_t\end{aligned}$$

where $\omega_t \sim i.i.d.\ N(0, \sigma_\omega^2)$, i.e., by allowing the slope of the trend function to follow a random walk with Normal errors. With the real GDP series, the estimate of σ_ω^2 is very small and this generalization leaves the trend-cycle decompositions virtually unchanged.

This analysis implies the following, provided the basic structure of the model (1) is adequate: 1) the trend dominates the series leaving only a small role to the cycle; 2) shocks to the trend are very negatively correlated with shocks to the cycle; and 3) the cycle bears no resemblance to the NBER chronology. Various explanations have been advanced to explain these results. For item (1), the common explanation is that technology shocks are mostly responsible for movements in aggregate production. These having a permanent effect, variations in real GDP show up as variations in the trend function. Hence, this type of results tends to support a real business cycle approach to movements in production leaving little room for monetary type explanations, which could account for the cyclical component. Explanations for item (2) follows from this real business cycle approach. For example, a positive shock to technology may imply that some labor skills become obsolete, thereby inducing a decrease in employment of a temporary nature until re-training is completed (note, however, that this type of explanation is harder to justify in the case of a negative technological shock). Finally explanations for item (3) often center on a distinction between "growth cycles" and "business cycles" (e.g., Zarnowitz and Ozyildirim, 2002). The NBER chronology is then viewed as pertaining to "business cycles" while decompositions of the type considered here pertains to document "growth cycles". Whatever the appeal, or lack thereof, of such explanations, the issue is ultimately an empirical one. It is therefore important to carefully assess whether the basic model (1) is free of important mispecifications.

Our argument will be that the basic model suffers from an important mispecification, which completely biases the results and their implications. A glimpse of our explanation can be gleaned from Figure 1, where it is seen that the cyclical component of the UC0 decomposition shows a marked decrease in mean from the pre to the post 1973 periods. The decrease in mean is such that the cyclical component completely misses the boom period of the late 90s and classify it as one of below trend activity. A more precise characterization of this feature can be obtained looking at the mean of the estimates of the residuals of the trend function η_t, using the filtered values. For the period pre-1973:1, the sample averages are 0.159 and 0.148 for the UCUR and UC0 decompositions, respectively, while for the post 1973:1 period the corresponding sample averages are -0.168 and -0.143. This is an economically important difference since it suggests a mean growth rate of the trend that is 1.31% (on an annual basis) lower after 1973:1 using the UCUR decomposition. The implied decrease is

1.16% using the UC0 decomposition, a figure that is smaller due to the fact that the change in the cyclical component also accounts for a part of the decrease (unlike the UCUR cycle which shows no apparent change). Given that the full sample estimate of the rate of growth μ is 3.24% on an annual basis (see Tables 1 and 3, Morley et al., 2003), the shocks to the trend function accounts for a 40% decrease in the overall rate of growth after 1973:1 (using the UCUR decomposition).

Hence, if we take the results suggested by Morley et al. (2003) and others at face values we are led to conclude that, on average, the post 1973 period has been subject to a sequence of negative shocks and the pre-1973 period enjoyed a sequence of positive shocks. While one may find appealing some ex-post justifications for the fact that the decomposition leaves little to the cycle, that the shocks to the trend and the cycle are negatively correlated and that the cycle bears no resemblance to the NBER chronology, it is hard to find any plausible explanation for sequences of shocks having different means and signs for the pre and post 1973 periods. Our aim is to show that all these seemingly puzzling results are artifacts of a neglected change in the slope of the trend function in 1973, and that once this is accounted for, all methods agree on a single decomposition, which is albeit very different.

3 Decompositions allowing a change in the trend function.

Our approach is to allow for the possibility of a permanent change in the trend function of real GDP occurring in 1973:1. To that effect, we introduce a simple modification to the basic model (1) such that

$$\tau_t = \mu + d1(t > T_b) + \tau_{t-1} + \eta_t$$

where $1(A)$ is the indicator function for the event A, and T_b is the observation corresponding to the time of break, 1973:1. The rest of the model stays the same. In particular, we shall continue to adopt an $AR(2)$ specification for the cyclical component.

It is important to discuss the implications of this, seemingly, minor change. First, we model the change in the trend function by a change in the deterministic component of the trend. This is to capture the fact that the change is viewed as a “permanent” one time change in the rate of growth. By “permanent”, we mean that the change is still in effect at the end of the sample period under consideration. Also, it is modelled as exogenously given to separate this change from the noise component. We shall return in Section 5 with a specification that allows for a stochastic change occurring at an unknown date. We start with this simple exogenous change occurring at a known date to better highlight how such a generalization leads to dramatically different results, quantitatively and qualitatively. Finally, note that our specification nests the previous ones. We are not forcing a change in the average rate of growth in 1973, we are simply allowing it to happen.

As a matter of notation, we denote the corresponding models by UC073, UCUR73 and ARIMA73. Technical details about the estimation are in Appendix B. The results for the parameter estimates are presented in Table 1 and the trend-cycle decompositions in Figure 2. The results are now strikingly different.

First, all three models agree on the point estimates of the rates of growth. It is 3.8% on an annual basis for the pre-1973 period and 2.64% for the post-1973 period, thereby indicating a 31% decrease. The UC073 specification, which constrains the shocks to the trend and cycle to be uncorrelated, shows a point estimate of $\sigma_\eta = 0$, implying a deterministic trend function, i.e., all random variations are captured by the cyclical component. For the UCUR73 specification, the point estimates are different; the standard deviation σ_η is slightly higher at 0.011 and the correlation of the shocks to the trend and the cycle is 1.0, perfectly positively correlated. The value of the likelihood function is, however, barely higher by a value of 0.19, for the UCUR73 model compared to the UC073. This suggests that the numerical results for the UCUR73 model may be due to an ill-behaved likelihood function or some in-sample overfitting problems. The numerical values for the ARIMA73 specification are slightly different, especially with respect to the AR coefficients for the cycle, and the likelihood fonction is higher by a value 1.58, which suggests that neither constraints imposed by the UC073 and UCUR73 specifications are exactly satisfied. Yet the point estimates of the moving-average coefficients sum to -1, which again indicates a deterministic trend function since the first-differences of real GDP is over-differenced.

Despite the numerical differences in the point estimates reported above, the trend-cycle decompositions for the 3 specifications, reported in Figure 2, are virtually identical. They clearly show a trend function that is piecewise linear (except at the very beginning of the sample period), with a clear decrease in the average rate of growth. The implied cycle is very different from those without the change in trend. It is important in magnitude and shows movements that corresponds very closely to the NBER chronology. Indeed, with two minor exceptions, a crossing of the zero axis from above occurs during a period identified by the NBER as a recession (the main exception is the recession of 1958, which would have been called a few quarters earlier according to our cycle). Also, unlike most trend-cycle decompositions proposed, ours clearly identifies the late 90s as a period of above-trend activity.

To summarize, the main qualitative features of the trend-cycle decompositions allowing for a change in the rate of growth in 1973:1 are: 1) all three specifications leads to the same conclusions, there are no conflicts anymore; 2) the average rate of growth has decreased 31% after 1973; 3) the trend function is piece-wise linear so that random variations are ascribes solely to the cyclical component; 4) the correlation of the shocks to the trend and cycle is trivially zero since the former is non-stochastic; 5) the cycle shows movements that follow closely the NBER chronology.

Note that these result are consistent with the findings of Perron (1989) who argued that once allowance is made for the possibility of a change in the trend function of real GDP, the null hypothesis of a unit root can be rejected. This result was criticized by several authors (e.g., Christiano, 1992, and Zivot and Andrews, 1992) who argued that one must take into account the possibility of data-mining induced by the ex-post choice of the break date and suggested method to treat the break date as unknown, in which case the unit root could no longer be rejected. Some of the ensuing literature viewed their result as convincing evidence that a unit root was present. Such a conclusion simply misses the fact that a failure to reject does not provide evidence about the null hypothesis, indeed the differences obtained may simply be due to a reduction in the power of the tests.

3.1 Additional Evidence

The fact that our ARIMA73 model shows a non-invertible moving-average structure and that our UC073 model has a point estimate of zero variance for the trend needs to be carefully assessed. The problem is that such estimates are likely to occur with some probability even if the true value is different. This is often referred to as the "pile-up" problem. In the case of the ARIMA73 specification, the problem is that if the sum of the moving-average coefficients is negative and close to -1, the probability distribution of the maximum likelihood estimate (MLE) of this sum will show a mass at the value -1. Similarly in the UC073 model, if the value of σ_η^2 is small, the MLE will also have a probability distribution with a mass at 0. So care must be exercised to assess the extent to which such a problem may be present.

If, as we conjecture, the UC073 or ARIMA73 specifications are appropriate ones for the data, two equivalent representations are an ARIMA(2,1,1) with a moving average coefficient of -1 or a trend-stationary model in levels of the form

$$\phi(L)(y_t - c - \mu t - d1(t > T_b)\,(t - T_b)) = e_t$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2$. The maximum likelihood estimates of these two models are presented in Table 2. For the ARIMA(2,1,1), the moving-average coefficient is indeed -1 and the estimates of the other parameters are virtually identical to those obtained from the UC073 specification. Similarly, the estimates of the model in level, which are unaffected by the pile-up problem are, again, basically identical. Hence, this is additional evidence pointing to the fact that the appropriate specification for the data is that delivered by the UC073 specification (the UCUR and ARIMA73(2,1,2) may be subject to inaccuracies resulting from an ill-behaved likelihood function even though the trend-cycle decomposition are similar). Note also that the AR(2) specification for the cyclical component is supported by the data. Both coefficients are significant and the Ljung-Box statistics applied to the estimated residuals $\hat{e}_t$ show no evidence of remaining serial correlation.

An alternative way to provide additional evidence for our proposed specification is to obtain a median unbiased estimate of σ_η^2 for the UC073 model. A method to do so was provided by Stock and Watson (1998). To implement this procedure, we write the UC073 model can as

$$\begin{aligned} y_t - c - \mu t - d1\left(t > T_b\right)\left(t - T_b\right) &= \tau_t + u_t \\ \tau_t &= \tau_{t-1} + \left(\frac{\lambda}{T}\right)\eta_t^* \\ \phi\left(L\right)u_t &= e_t \end{aligned} \tag{2}$$

or in first-differences, as

$$\Delta y_t - \mu - dI_{(t>T_b)} = \frac{\lambda}{T}\eta_t^* + \Delta u_t$$

where $\eta_t^* \sim i.i.d.\ N(0,1)$ so that $\eta_t = (\lambda/T)\eta_t^* \sim i.i.d.\ N(0,(\lambda/T)^2)$. This specifies that the variance of the trend function is "close to" zero, pertaining to cases where the pile-up problem may occur. Stock and Watson (1998) provide methods to construct a median-unbiased estimate of λ as well as a confidence interval. The method relies on the fact that the component $\tau_t + u_t$ will show structural changes in levels if $\lambda \neq 0$, the extent of which depends on the parameter λ. The idea is then to apply a structural change test to this component and back out from it a confidence interval and the median unbiased estimate of λ. Since $\tau_t + u_t$ is unobserved, one applies the procedure to the least-squares residuals from a regression of y_t on a constant, and the split trend. Stock and Watson (1998) suggest a variety of structural change tests to perform this procedure. We applied all of them and the results are presented in Appendix C. The results are unanimous. The median unbiased estimate of λ is 0, there is no evidence of a pile-up problem (sensitivity analyses showed that the same results obtain with alternative specifications for the order of the autoregressive process for the cyclical component). This reinforces our conclusion that the appropriate specification is that provided by the UC073 model.

4 Can the UC073 model explain previous results?

The next issue we wish to address is the following: assuming that the true data generating process is one with a piecewise linear trend with zero variance and an AR(2) noise component, can we explain the estimates found in Section 2 based on specifications that do not allow a change in the rate of growth in 1973? To answer this question, we resort to simulation experiments. The data is generated by the following broken-trend stationary process

$$\begin{aligned} y_t &= a + \mu t + d 1 \left(t > T_b\right) \left(t - T_b\right) + c_t \\ c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + e_t \\ e_t &\sim i.i.d.\ N\left(0, \sigma_e^2\right) \end{aligned}$$

with the following parameters: $a = 724.18$, $\mu = 0.95$, $\phi_1 = 1.38$, $\phi_2 = -0.28$ and $\sigma_e^2 = 0.94$. The sample size is set to $T = 200$ and the break is assumed to occur at mid sample, i.e. $T_b = 100$ (we generate 206 value and discard the first 6 observations). The base case considers a change in slope $d = -0.29$, consistent with the estimate obtained. However, we also consider simulations with $d = -0.1, -0.4$ and -0.6 to better assess the effect of the change in trend on the key parameters of interest. For each set of generated data, we compute the parameter estimates of the UC0, UCUR and ARIMA(2,1,2) models. For the UC0 and UCUR models, we also compute the means of the median values of the filtered estimates of the residuals of the trend function, η_t. This is repeated until we have 200 draws for which the estimation was successful in the sense that convergence was achieved. The computer language used in this simulation is MATLAB, and the maximization was implemented using the command '`fminunc`'. For the initial values, we used the output value of the original estimation by Morley et al. (2003). When evaluating the likelihood function, if any matrix becomes close to singular, we skip that replication and treat it as having "not converged". The condition used is that the inverse of the condition number is less than machine epsilon ('`rcond_(X)<eps`').

The results for the base case are presented in Table 3, which for convenience also reproduces the estimates reported in Tables 1-3 of Morley et al. (2003). Consider first the UC0 specification. The simulated values obtained are indeed very close to the sample values, certainly within Monte Carlo standard errors. For the UCUR and ARIMA models the match is not as good but the values obtained are again within Monte Carlo standard errors of the sample estimates. But the key point is that the simulated values reveal that the piece-wise trend stationary structure can explain the main qualitative findings that occur when the change in slope in ignored. First, the variance of the residual to the trend function is biased away from zero suggesting a stochastic trend. Secondly, in the UCUR model, the correlation of the shocks to the trend and the cycle is large and negative.

Table 4 shows the same simulation exercise for different values of the change in slope d. It is clear from the results that, as the change in slope increases in absolute value, the variance of the shocks to the trend, σ_η^2, increases, and, for the UCUR model, the correlation $\rho_{\eta e}$ approaches -1. Hence, these features, which may be puzzling to some, can easily be accounted for by the neglected change in slope.

But why is this the case? The fact that σ_η^2 is biased away from 0 in the presence of a change in slope that is unaccounted for is simply a manifestation of the phenomenon documented by Perron (1989), namely that a trend-stationary process with a change in slope can appear as a unit root process. The fact that the correlation is biased towards -1 follows basically as a Corollary. Indeed, since σ_η^2 is inflated, large fluctuations are ascribed to the trend. Hence, to compensate, cyclical fluctuations in the opposite direction are needed.

Finally, Table 5 presents the mean of the average values, for the pre (using observations from $t = 3$ onwards) and post break samples, of the median (over the 200 replications) of the filtered estimates of the residuals of the trend function. The results confirm the following features: 1) for both the UC0 and UCUR models, the mean is positive for the pre-break period and negative for the post-break period; 2) the difference is bigger for the UCUR model than for the UC0 model (recall that the cyclical component also shows a change in level for the UC0 model); 3) the spread increases as the magnitude of the change increases; 4) the base case with $d = -0.29$ delivers values quite close to the sample estimates.

5 An alternative framework for trend-cycle decompositions

Our aim in this section is to suggest a class of unobserved components models that is able to capture structural changes in the trend function endogenously. While we believe our suggested framework to be powerful, we make no claim about its optimality, nor do we claim to have solved all problems related to identification and numerical estimation. Nevertheless, we will show how it can be a powerful tool for issues related to trend-cycle decompositions and, hence, a viable avenue for further developments.

To understand the need for the key ingredients, let us go back to the generalized trend function

$$\begin{aligned} \tau_t &= \beta_t + \tau_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + v_t \end{aligned} \tag{3}$$

As stated in Remark 1, this specification provides very similar results compared to the case where β_t is assumed fixed, when the noise to the slope component β_t is assumed $i.i.d.$ $N(0, \sigma_v^2)$. The reason is that any positive variance σ_v^2 would imply changes in the slope occurring at every period, though of different magnitudes each time. Now, if our proposed specification is adequate, the slope of the trend function changes very rarely, indeed it is expected to change only once, or if the change occurs smoothly, for a few periods around 1973:1. This is the key observation and it suggests the use of a non-Normal distribution for the errors v_t. The natural one to adopt in our context is a mixture of Normal distribution where a realization of v_t is a draw from one of two Normal distributions, one with high and the other with small or zero variance. More specifically,

$$v_t = \lambda_t \gamma_{1t} + (1 - \lambda_t)\, \gamma_{2t} \tag{4}$$

where $\gamma_{it} \sim i.i.d.\ N(0, \sigma_{\gamma i}^2)$ and λ_t is a Bernoulli random variable that takes value one with probability α_1, and value 0 with probability $1 - \alpha_1$. In our case, we would expect α_1 to be close to one and $\sigma_{\gamma 1}^2$ to be zero, so that most of the time there is no change in the slope of the trend function. Furthermore, if $\sigma_{\gamma 2}^2 > 0$, there will be occasional changes to the value of the slope. Hence, this specification appears ideally suited to the problems we face.

More generally, the unobserved component model we propose is the following:

$$y_t = \tau_t + c_t + \omega_t \tag{5}$$

where τ_t is specified by (3) with $\eta_t \sim i.i.d.\ N(0, \sigma_\eta^2)$ and v_t is generated by the mixture distribution (4). The component ω_t is introduced to capture measurement errors and is assumed to be $i.i.d.\ N(0, \sigma_\omega^2)$. The cyclical component will still be assumed to be an AR(2) but we shall also generalize it to have shocks generated by a mixture of Normals as well, i.e., we have

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \epsilon_t \tag{6}$$

where

$$\epsilon_t = \delta_t \xi_{1t} + (1 - \delta_t) \xi_{2t} \tag{7}$$

with $\xi_{it} \sim i.i.d.\ N(0, \sigma^2_{\xi i})$ and δ_t a Bernoulli random variable that takes value one with probability α_2, and value 0 with probability $1 - \alpha_2$. This generalization is made to potentially capture the fact that the variance of recessions may be different from the variance of expansions.

So the complete model consists of the specifications (3) to (7) with the added assumption that all errors terms and Bernoulli random variables are mutually independent.

This type of model is not new and has been used in the statistics literature to model structural changes; see, in particular, Kitagawa (1987). It is a generalized State Space model where some of the errors are non Normal. The State Space model is of the form

$$\begin{aligned} y_t &= Hx_t + \omega_t \\ x_t &= Fx_{t-1} + Gu_t \end{aligned}$$

where $x_t = [\tau_t, c_t, c_{t-1}, \beta_t]'$, $H = [1, 1, 0, 0]$

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

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and $u_t = [\eta_t, \varepsilon_t, v_t]'$. What is different from the usual State Space model is that the distribution of u_t is not Normal. However, we can view the specification as a State Space model with Normal errors but with four possible states. These states are defined by the combined values of the Bernoulli random variables λ_t and δ_t and imply four possible covariance matrices for the vector of errors u_t, namely

$$Q = \left\{ \begin{bmatrix} \sigma^2_\eta & 0 & 0 \\ 0 & \sigma^2_{\gamma 1} & 0 \\ 0 & 0 & \sigma^2_{\xi 1} \end{bmatrix}, \begin{bmatrix} \sigma^2_\eta & 0 & 0 \\ 0 & \sigma^2_{\gamma 1} & 0 \\ 0 & 0 & \sigma^2_{\xi 2} \end{bmatrix}, \begin{bmatrix} \sigma^2_\eta & 0 & 0 \\ 0 & \sigma^2_{\gamma 2} & 0 \\ 0 & 0 & \sigma^2_{\xi 1} \end{bmatrix}, \begin{bmatrix} \sigma^2_\eta & 0 & 0 \\ 0 & \sigma^2_{\gamma 2} & 0 \\ 0 & 0 & \sigma^2_{\xi 2} \end{bmatrix} \right\}$$

where each component occurs with probabilities $\alpha_1\alpha_2$, $\alpha_1(1-\alpha_2)$, $(1-\alpha_1)\alpha_2$, and $(1-\alpha_1)(1-\alpha_2)$, respectively. This interpretation is helpful in constructing an algorithm for estimation. Note also that this specification is different from the popular Markov type regime switching model (e.g., Hamilton, 1989). Our structure does not assume Markov type transition probabilities but rather that each state is drawn independently with some probability.

Our generalization complicates the estimation procedure considerably and details are given in Appendix C. The basic principles are, however, the same as for the estimation of the usual State Space model with Normal errors. The likelihood function is estimated using a variant of the Kalman filter and a by-product is an estimate of the conditional expectation of the State vector x_t using information available up to time t. These are denoted $x_{t|t}$ and are called filtered estimates. One can also construct estimates using the full sample, i.e., $x_{t|T}$ which are obtained using a smoothing algorithm and are, accordingly, called smoothed estimates. The main goal here is to obtain smoothed estimate of the trend function τ_t and of the cyclical component c_t.

It is important to note that, as stated, not all parameters are identified, though the trend-cycle decomposition is. To get parameter estimates we impose the following restrictions: $\alpha_1 > 0.9$, $\sigma^2_{\gamma 1} < 0.0001$, $\sigma^2_{\gamma 1} < \sigma^2_{\gamma 2}$ and $\sigma^2_{\xi 1} < \sigma^2_{\xi 2}$. The last two restrictions are standard and inconsequential. The first two are, however, more substantive. They impose the variance of the state occuring with highest probability be very small and that the latter probability be quite high. This is to allow for the possibility of having relatively rare events occuring to the trend function. Changes in these restrictions do no change the trend-cycle decomposition but do change the parameter estimates obtained. More work is needed to carefully assess the identification of such models. Meanwhile, since the main object of interest is the trend-cycle decomposition, we feel confident with the results since this decomposition is identified.

5.1 Discussion of the results

The results are presented in Table 6 and Figure 3. The most important element is the smoothed trend-cycle decomposition presented in the bottom panel of Figure 3. It shows a trend and a cycle that are qualitatively similar to those obtained imposing an exogenous break in the slope of the trend in 1973:1. To better understand its properties, it is useful to look at the parameters estimates presented in Table 6.

First, and most important, is the fact that the estimate of the variance of the residuals of the trend function, σ^2_η is estimated to be 0. Hence, except for changes in the slope β_t, the trend is deterministic. The innovations to the trend function are governed by a process that has standard deviation .0001 with probability .9 and one that has standard deviation 0.0633 with probability .10. These estimates are, however, highly dependent on the restrictions imposed and we shall return below with a better method to identify the pattern of the slope of the trend function. But the basic message is clear, most of the time the slope does not vary. A look at Figure 3 suggests that the changes occur smoothly around 1973:1.

Other estimates are also very informative. The noise of the cyclical component also consists of draws from Normal distributions with very different variances. With probability .58 the standard deviation is small at the value 0.26, and with probability .42 it is high at the value 1.28. We interpret these results are follows. The variance of shocks in recessions are much larger than the variance of shocks in expansions. A look at the smoothed cycle in Figure 3 shows this to be indeed the case. Recessions are much more pronounced than expansions. This accords well with previous studies on business cycles asymmetries; see, e.g. Beaudry and Koop (1993), Neftci (1984) and Sichel (1993, 1994). The parameter estimates of the AR coefficients are well within the stationary region, the sum being close to 0.91. The variance of the measurement errors is quite small and does not account for much of the movements of real GDP. Finally, note that the value of the maximized likelihood function is -272.5, well above that for the models discussed earlier.

The smoothed estimate of the cycle share many of the interesting features that were present for the models with an exogenous change in 1973:1. First, the movements agree quite well with the NBER chronology. Second, the late 90s are, as should be expected, characterized by above trend activity. Third, as alluded to above, recessions are characterized by sharp drops in activity, while expansions are gradual increases. Fourth, the sharpest recession is that of 1982, while our model with an exogenous change in slope and symmetric errors for the cyclical component indicated the recession of 1958 as the sharpest. Fifth, depth of recessions are larger than highs of expansions in the sense that recessions are often characterized by a value 6 to 8 percentage points below trend, while expansions are characterized by values that reach between 2 and 4 percentage points above trend.

The filtered estimates of the trend and cycles also show some interesting features. This filtered decomposition gives the best estimates of the trend and cycle using only the information available up to the current period. One fact is that the slope of the trend seemed to be on the increase in the late 1990s. This accord well with discussions at the time that the trend may have been on a new path with the new information technology. However, a comparison with the smoothed trend, which uses all information in the sample, shows this hope not to have materialized (at least by the end of 1998). Another feature of interest is that the sharpness of the 1991 recession was ex-post more severe than what could have been inferred at the time.

We believe these results to be in accordance with common notions about the business cycles and to lend credence to our framework as a general methodology for trend-cycle decomposition.

5.2 Filtered estimates for the slope of the trend function

As discussed, while the trend-cycle decomposition is well identified, the identification of some components is more problematic. Since the temporal behavior of the slope of the trend function is of central concern here, we present a two-step method to document it. We start with the filtered estimates of the trend function, $\{\widetilde{\tau}_t = \tau_{t|t}\}$, as the basic inputs. We then estimate the following model

$$\begin{aligned} \widetilde{\tau}_t &= \beta_t + \widetilde{\tau}_{t-1} + \widetilde{\eta}_t \\ \beta_t &= \beta_{t-1} + v_t \end{aligned} \tag{8}$$

where

$$\widetilde{\eta}_t = \left(\tau_t - \tau_{t|t}\right) - \left(\tau_{t-1} - \tau_{t-1|t-1}\right) + \eta_t.$$

and v_t is specified by the mixture of Normal distribution (4). This model is simple enough that an exact numerical procedure to obtain smoothed estimates of the slope is possible, i.e., $\beta_{t|T}$. The Fortran algorithm was constructed by Kitagawa (1993) and we translated it in Matlab for the estimation reported (this is useful since such an exact algorithm avoids the approximations that are necessary to estimate the full model, see Appendix C).

The results are presented in Figure 4. The solid line is the smoothed estimate of the slope $\beta_{t|T}$ and the dashed lines are 1, 2 and 3 standard deviations intervals. The results are quite informative and in line with our main argument. The slope starts at a value 0.94 at the beginning of the sample and remains at that value until roughly the late 60's when a minor decrease starts to take effect. The main change occurs in the period 1973-1974 when the slope decreases to a value of approximately 0.75. Until 1977 there is a further gradual decline to a new value of 0.66 that remains in effect until the end of the sample.

These results lend support to the central them of our work, namely that an important change in the slope of the trend function has occurred around the year 1973. While, the change depicted here is more gradual than the assumed sudden change used in the previous sections, the message is the same. The change is important and is responsible for the severe biases arising when estimating models that neglect its presence.

6 Comparisons with other decompositions

Many trend-cycle decompositions have been suggested in the literature. Besides the Beveridge-Nelson decomposition and the Unobserved Components models examined previously, two other popular methods are the Hodrick-Prescott (1997), or HP, filter and the Band Pass, or BP, filter (Baxter and King, 1999). The former is a method to extract a trend function and delivers, as a consequence, a cycle as the difference with the original series. The BP filter, however, does not address the issue of trend estimation. The cycle is defined as movements having periods between 6 and 32 quarters. Hence, both high and low frequency movements are eliminated and the difference between the cycle obtained and the original series cannot be viewed as an estimate of the trend function.

In this section, we compare how these trend and cycle extraction procedures compare with our decomposition. Figure 5 presents the cycle obtained from the HP filter with the usual value of the smoothing parameter $\lambda = 1,600$. Figure 6 presents the cycle obtained from the BP filter using the two-sided filter suggested by Baxter and King (1999) with 12 terms on each sides (accordingly the cycle is undefined for the first and last 12 quarters of the sample).

The cycles obtained are somewhat "in between" our decomposition and the BN cycle advocated by Morley et al. (2003). Both show much less variations than our cycle and slightly more than the BN cycle. Also, the movements are more frequent than in our cycle but less so than in the BN cycle. Overall, they capture rather well the timing of the recessions (as does our cycle) but the depth of the recessions and the heights of the expansions are very different. For example, both the HP and BP cycles show the sixties as a period of average activities or very mild expansion following a mild recession in 1961, while ours characterize the sixties as a period of important and sustained expansion following a deep recession in 1961. Other differences can be ascertained from the graphs, but the most striking feature is the magnitude of the cycle. The HP trend being stochastic, a lot of the movements in real GDP are due to movements in the trend, while little is left for the cycle. While the BP filter does not estimate a trend directly, the overall picture is much the same.

Finally, Figure 7 presents an interesting result about the HP filter. It gives the trend-cycle decomposition when the smoothing parameter is set to the very large number $\lambda = 800,000$. The results are then basically equivalent to our decomposition with a trend that has a single shift in slope in 1973. So we can, in a sense, reconcile our results with the decomposition of the HP filter. The latter depends crucially on the choice of the smoothing parameter λ and, as is well known, the best value depends on the underlying true structure of the process and the usual rule of thumb of setting $\lambda = 1,600$ is unlikely to be appropriate in most cases. Our decomposition does not suffer from such arbitrariness and suggests that for the particular case of the real GDP series analyzed, $\lambda = 1,600$ is too small a value that has an effect of ascribing too much variation to the trend and too little to the cycle.

7 Concluding Remarks

The title of the paper by Morley et al. (2003) asks the following question: "Why are the Beveridge-Nelson and Unobserved-Components Decomposition of GNP so Different?". Their answer was that the latter does not allow for the presence of negative correlation between the noise to the trend and the noise to the cycle and is, hence, mispecified. They also argued that an unobserved component model that allows for this feature is consistent with the data and yields a decomposition virtually identical to the Beveridge-Nelson decomposition. Our results, based on allowing for a one time change in the slope of the trend function in 1973, offers a rather different picture. First, the question is ill-posed. Both methods yields the same decomposition. Second the trend function is non-stochastic except for a brief period around 1973 and, hence, their is no issue about the correlation between the noise to the trend and the cycle, they are uncorrelated.

Our results also show the trend-cycle decomposition to be very different from that obtained using currently popular methods (BN or UC decompositions, HP or BP filters). It also agrees much better with the NBER chronology and has the advantage of being able to explain results pertaining to decompositions obtained by methods such as that of Beveridge and Nelson and the Unobserved Components models.

We also presented a generalized Unobserved Components model based on errors that follow a mixtures of Normal distribution. It was found to be successful in reaching the same conclusions without the need to make any prior specifications about the nature or timing of the change in the slope. Such a framework should find wide appeal for trend-cycle decompositions in a variety of contexts, when the trend path of some series is affected by infrequent level shifts or changes in slope, or when the noise function shows different variability across two regimes, or when the overall series is affected by aberrant observations.

Needless to say, our work is a first attempt in using such Non-Gaussian State Space models with this level of generality, at least in economics, and there is certainly a need to refine the framework and estimation method used. One useful generalization would be to allow for the possibility of correlation across the stochastic components of the trend and the cycle. While we have documented that such a feature is not present with the particular US real GDP series used, this may not be the case in general. Also, our experience has shown that the estimation of such models is a difficult task, mainly due to a likelihood function that is not well behaved. While all our trials showed the trend-cycle decomposition to be well identified, this is not so for some subsets of the parameters and makes difficult a careful interpretation of the results. Developments to solve or mitigate these problems are important. Finally, while most decompositions used in applied work is based on univariate methods, extensions to decompositions based on modelling multiple series jointly are important.

References

[1] Alspach, D.L. and H.W. Sorenson (1972) "Nonlinear Bayesian Estimation using Gaussian Sum Approximation," *IEEE Transactions on Automatic Control* 17, 439-448.

[2] Andrews, D.W.K. and W. Ploberger (1994) "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica* 62, 1383-1414.

[3] Ansley, C.F. and P. Newbold (1980) "Finite Sample Properties of Estimators for Autoregressive Moving Average Models," *Journal of Econometrics* 13, 159-183.

[4] Baxter, M. and R. King (1999) "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics* 81, 575-593.

[5] Beaudry, P. and G. Koop (1993) "Do Recessions Permanently Change Output?," *Journal of Monetary Economics* 31, 149-163.

[6] Beveridge, S. and C.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, 151-74.

[7] Burns, A.M. and W.C. Mitchell (1946) *Measuring Business Cycles*, National Bureau of Economic Research: New York.

[8] Campbell, J. and N.G. Mankiw (1987) "Are Output Fluctuations Transitory?," *Quarterly Journal of Economics* 102, 857-880.

[9] Canova, F. (1998) "Detrending and Business Cycle Facts," *Journal of Monetary Economics* 41, 475-512.

[10] Canova, F. (1999) "Does Detrending Matter for the Determination of the Reference Cycle and the Selection of Turning Points,' *Economic Journal* 109, 126-150.

[11] Clark, P.K. (1987) "The Cycle Component of the U.S. Economic Activity," *Quarterly Journal of Economics* 102, 797-814.

[12] Cochrane, J. (1988) "How Big Is The Random Walk in GNP?," *Journal of Political Economy* 96, 893-920.

[13] Durbin, J. and S.J. Koopman (2000) *Time Series Analysis by State Space Methods,* Oxford University Press: New York

[14] Hamilton, J.D. (1989) "A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycles," *Econometrica* 57, 357-384.

[15] Harrison, P.J. and C.F. Stevens (1976) "Bayesian Forecasting," *Journal of the Royal Statistical Society*, Series B, 38, 205-247.

[16] Hodrick, R. and E. Prescott (1997) "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking* 29, 1-16.

[17] Kim, C.-J. (1994) "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics* 60, 1-22.

[18] Kim, C.-J. and C. Nelson (1999) *State Space Models with Regime Switching,* MIT Press: Cambridge.

[19] Kitagawa, G. (1987) "Non-Gaussian State-Space Modeling of Nonstationary Time Series," *Journal of American Statistic Association*, 82, 1032-1063.

[20] Kitagawa, G. (1993) *FORTRAN77 Time Series Analysis Programming*, Iwanami-Shoten: Tokyo (in Japanese).

[21] Morley, J.C. (2000) "A State-Space Approach to Calculating the Beveridge Nelson Decomposition," *Economics Letters* 75, 123-127.

[22] Morley, J.C., C.R. Nelson and E. Zivot (2003) "Why are Beveridge-Nelson and Unobserved-Component Decompositions of GDP So Different?," *Review of Economics and Statistics* 85, 235-243.

[23] Neftci, S. N. (1984) "Are Economic Time Series Asymmetric Over the Business Cycles?," *Journal of Political Economy* 92, 307-328.

[24] Nyblom, J. (1989) "Testing for the Constancy of Parameters over Time," *Journal of the American Statistical Association* 84, 223-230.

[25] Perron, P. (1989) "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica* 57, 1361-1401.

[26] Quandt, R.E. (1960) "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association* 55, 324-330.

[27] Saikkonen, P. and R. Luukonen (1993) "Point Optimal Test for Testing the Order of Differecing in ARIMA Models," *Econometric Theory* 9, 343-362.

[28] Sargan, J.D. and A. Bhargava (1983) "Maximum Likelihood Estimation of Regression Models with First Order Moving Average Errors with the Root Lies on the Unit Circle," *Econometrica* 51, 799-820.

[29] Shively, T.S. (1988) "An Exact Test of a Stochastic Coefficient in a Time Series Reegression Model," *Journal of Time Series Analysis* 9, 81-88.

[30] Sichel, D.E. (1993) "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry*, 224-236.

[31] Sichel, D.E. (1994) "Inventories and the Three Phases of the Business Cycle," *Journal of Business and Economic Statistics* 12, 269-277.

[32] Stock, J.H. and M.W. Watson (1988) “Variable Trends in Economic Time Series,” *Journal of Economic Perspectives* 2 (3), 147-174.

[33] Stock, J.H. and M.W. Watson (1998) “Median Unbiased Estimation of Coefficient Variance in Time-Varying Parameter Model,” *Journal of American Statistical Association* 93, 349-358.

[34] Stock, J.H. and M.W. Watson (1999) “Business Cycle Fluctuations in US Macroeconomic Time Series," in J.B. Taylor and M. Woodford (eds.) *Handbook of Macroeconomics*, vol. 1A, Amsterdam: North-Holland.

[35] West, M. and J. Harrison (1997) *Bayesian Forecasting and Dynamic Models,* Springer Verlag: New York.

[36] Watson, M. (1986) “Univariate Detrending Methods with Stochastic Trend,” *Journal of Monetary Economics* 18, 49-75.

[37] Zarnowitz, V. and A. Ozyildirim (2002) “Time Series Decompositions and Measurement of Business Cycles, Trends and Growth Cycles," NBER Working Paper 8736, National Bureau of Economic Research: Cambridge.

A-1 Appendix A: Details on the computations in Section 2.

In Section 2 to 4, unless otherwise indicated, the estimation of the models is done by casting them in the appropriate State Space model and estimating their parameters by Maximum Likelihood using the Kalman Filter algorithm. In general, the State Space model is of the form

$$
\begin{aligned}
y_t &= Hx_t && \text{(A.1)}\\
x_t &= M + Fx_{t-1} + Gu_t && \text{(A.2)}\\
E(u_t u_t') &= Q
\end{aligned}
$$

where (A.1) is the measurement equation and (A.2) is the transition equation. For the UCUR model, $x_t = [\tau_t, c_t, c_{t-1}]'$, $u_t = [\eta_t, \varepsilon_t]'$, $H = [1, 1, 0]'$, $M = [\mu, 0, 0]'$

$$
F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad Q = \begin{bmatrix} \sigma_\eta^2 & \sigma_{\eta\varepsilon} \\ \sigma_{\eta\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix}
$$

The UC0 model is obtained as a special case with $\sigma_{\eta\varepsilon} = 0$. For the ARIMA(2,1,2) the measurement equation is

$$
\Delta y_t - \mu = Hx_t
$$

with $H = [1, 0, 0, 0]'$ and the state vactor is of dimension 4 with transition equation given by (A.2) with $M = 0$ and

$$
F = \begin{bmatrix} \phi_1 & \phi_2 & \theta_1 & \theta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}
$$

$G = [1, 0, 1, 0]'$, $u_t = e_t$ and $Q = \sigma_e^2$.

A-1.1 Numerical Estimation

For the numerical estimation, we used both a Gauss code provided by Morley written in the Gauss language and based on the 'optmum' command. We also estimated all models using an independently constructed code in Matlab 6.5 using the command 'fminunc'. Both are based on a quasi-Newton method and the numerical Hessian is updated by the BFGS equation.

For the Gauss code the tolerance for the gradient change is less than $1e-5$. For the Matlab code the stopping rule is slightly different and is based on i) a tolerance for parameter change less than $1e-6$ (`TolX`); and ii) a tolerance for a value function change of $1e-6$ (`TolFun`).

The restrictions for the parameters are imposed by reparameterizations so that i) the covariance matrix is positive definite and ii) the AR coefficients satisfy the stationarity conditions. Throughout, the computation is skipped and the results not used when a reliable inverse for a matrix cannot be obtained, i.e., when the condition number of the matrix is less than the machine epsilon (in our case, $2.2204e-16$).

As initial conditions, $x_{0|0}$ and $P_{0|0}$, we use the steady state values as in Morley et al. (2003). For a nonstationary component, we use a very large value on the corresponding diagonal element of $P_{0|0}$ and the first observations into the first row of $x_{0|0}$. The likelihood function is computed using the prediction errors from time $t=2$ onward.

A-2 Appendix B: Details on the computations in Section 3.

The estimation method is the same except that for all models μ is replace by $\mu_t = \mu + d1(t > T_b)$ where T_b is the observation corresponding to 1973:1. The level estimation is also implemented by maximizing the likelihood function of the State Space model specified by:

$$\Delta y_t - c - \mu t - d1(t > T_b)\,(t - T_b) = Hx_t$$

with $H = G = [1, 0]'$, $u_t = e_t$, $Q = \sigma_e^2$ and

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

The results were also confirmed using the Gauss `arima` routine.

To maximize the chances of obtaining values corresponding to the global maximum of the likelihood function, we estimated each model 1000 times using randomly generated initial values. Since we use a reparameterization such that Q is ensured to be positive definite and the cyclical component is stationary, the random initial conditions were generated as draws from a $N(0, 3)$. We then ranked the results according to the value of the likelihood function. For the UCUR (UC0, resp.) model, the first 200 (300, resp.) sets of results gave almost the same parameters. For the $ARIMA(2, 1, 2)$ model, the initial conditions for the MA coefficients are random draws from a $N(0, 1)$. More than 600 sets results were such that the sum of the MA coefficients was minus one, with almost the same AR coefficients. We performed this using codes in both GAUSS and MATLAB, which gave almost the same parameter estimates. The results reported in the Tables are those obtained with GAUSS. The reported numerical standard errors of the parameters are obtained using the Gauss commands '`hessp`' and '`gradfd`' in the prototype code provided by Morley. Yet, we failed to reproduce their results using our computer environment and were not able to get results simular to either using the Matlab code. Hence, these standard errors should not be considered reliable, and we do not make any inference based on them.

For the ARIMA models, we did not rely on the GAUSS `arima` library since its algorithm tries to avoid getting non-invertible MA roots, hence it sometimes converges a the local maximum in the strictly invertible region.

To check the adequacy of the AR(2) specification, we constructed the Ljung-Box test for the residuals of the level estimation. We cannot reject the null hypothesis that there is no serial correlation (the p-value is .28 with both 12 and 24 lags). The test statistic is constructed using the standardized forecast errors (for more details, see Durbin and Koopman, 2000).

A-2.1 Median Unbiased Estimate

To implement the median unbiased estimate of λ in the specification (2) we apply the procedures suggested by Stock and Watson (1998) to the residuals from a regression of log real GDP on a piece-wise linear trend with a change in slope in 1973:1. The idea of the procedure is as follows. Consider a test of the null hypothesis that $\lambda = 0$, say F_T. Since the model specifies a variance that is local to zero, such tests will have a non-degenerate limit distribution that will depend on the parameter λ such that, say, $F_T \Rightarrow F(\lambda)$. Consider a transformation $g(F(\lambda))$ and suppose there exists a monotone increasing function $m(\lambda)$, which gives the median of $g(F(\lambda))$. Then, the median unbiased estimate of λ is the value $\hat{\lambda}_{MU}$ such that $\hat{\lambda}_{MU} = m^{-1}(g(F_T))$. As in Stock and Watson (1998), the function $g(\cdot)$ is chosen to be the cumulative distribution function of $F(\lambda)$.

We applied this procedure using a variety of tests: Nyblom's (1989)'s L (L), the mean and exponential Wald tests of Andrews and Ploberger (1994) (MW, EW), Quandt's (1960) likelihood ratio test (QLR), the point optimal invariant test for $\lambda = 7$ ($POI7$), and for $\lambda = 17$ ($POI17$) of Saikkonen and Luukkonen (1993) and Shively (1988). All procedures were per-

formed using a GAUSS code available on Mark Watson's home page (http://www.wws.princeton.edu/~mwa
An AR(2) specification for the cyclical component was used. The results are presented in Table A.1.

Table A.1: Stock and Watson's (1998) Median Unbiased Estimate of λ

Test	Test Statistics	p value	λ	90% C.I.	$\sigma_{\Delta\beta}$	90% C.I.
L	0.0373	(0.9450)	0	$(0, 0.8440)$	0	$(0, 0.0403)$
MW	0.2323	(0.9050)	0	$(0, 3.4306)$	0	$(0, 0.1647)$
EW	0.1282	(0.9050)	0	$(0, 3.4507)$	0	$(0, 0.1646)$
QLR	1.3222	(0.9200)	0	$(0, 3.2556)$	0	$(0, 0.1553)$
$POI(7)$	1.5913	(0.8600)	0	$(0, 5.3004)$	0	$(0, 0.2529)$
$POI(17)$	6.8392	(0.6500)	0	$(0, 9.2410)$	0	$(0, 0.4409)$

The results are consistent across all tests: 1) the null hypothesis of a zero value of λ cannot be rejected, with p-values ranging from 0.65 to 0.95; 2) the median unbiased estimate of λ is zero using any test; 3) the confidence interval for λ and $\sigma_{\Delta\beta}$ is narrow and include the value 0. Hence, we view this as providing no evidence against the hypothesis that $\lambda = \sigma_{\Delta\beta} = 0$, or that the trend is deterministic. Also, sensitivity analyses revealed the same conclusions to hold for different specifications of the autoregressive order of the cyclical component.

A-3 Appendix C: Details on the computations of Section 5

Let $Y_t = (y_1, ..., y_t)$ be the vector of data available up to time t. The objective function to be maximized is

$$\ln(L) = \ln\left[\sum_{t=1}^{T} p\left(y_t|Y_{t-1}\right)\right]$$

$$p\left(y_t|Y_{t-1}\right) = \sum_{s_t=1}^{4}\sum_{s_{t-1}=1}^{4} p\left(y_t|s_{t-1}, s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i, s_t=j|Y_{t-1}\right)$$

Also let the prediction errors be

$$\nu_{t|t-1}^{ij} = y_t - E[y_t|Y_{t-1}, s_{t-1}=i, s_t=j] = y_t - Hx_{t|t-1}^{ij}.$$

Conditional on the states at periods t and $t-1$ taking values i and j, respectively, and the value of Y_{t-1}, the prediction errors are such that

$$(\nu^{ij}_{t|t-1}|s_{t-1}=i,s_t=j,Y_{t-1}) \sim N\left(0, f^{ij}_{t|t-1}\right) \tag{A.3}$$

with

$$f^{ij}_{t|t-1} = E\left(\nu^{ij}_{t|t-1}\nu^{ij\prime}_{t-1}\right) = HP^{ij}_{t|t-1}H' + G$$

so that

$$p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right) = \frac{1}{\sqrt{2\pi}}\left|f^{ij}_{t|t-1}\right|^{-1/2}\exp\left\{-\frac{\nu^{ij\prime}_{t|t-1}(f^{ij}_{t|t-1})^{-1}\nu^{ij}_{t-1}}{2}\right\}$$

Also,

$$\begin{aligned}
\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right) &= \Pr\left(s_t=j|s_{t-1}=i\right)\Pr\left(s_{t-1}=i|Y_{t-1}\right)\\
&= \Pr\left(s_t=j\right)\Pr\left(s_{t-1}=i|Y_{t-1}\right)\\
\Pr\left(s_{t-1}=i,s_t=j|Y_t\right) &= \Pr\left(s_{t-1}=i,s_t=j|y_t,Y_{t-1}\right) = \frac{p\left(y_t,s_t,s_{t-1}|Y_{t-1}\right)}{p\left(y_t \mid Y_{t-1}\right)}\\
&= \frac{p\left(y_t|s_t,s_{t-1,}Y_{t-1}\right)\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right)}{p\left(y_t|Y_{t-1}\right)}\\
\Pr\left(s_t=j|Y_t\right) &= \sum_{i=1}^{4}\Pr\left(s_{t-1}=i,s_t=j|Y_t\right).
\end{aligned}$$

The basic inputs are therefore the best estimates of the sate vector and their mean squared errors, namely

$$\begin{aligned}
x^{ij}_{t|t-1} &= Fx^{i}_{t-1|t-1}\\
P^{ij}_{t|t-1} &= FP^{i}_{t-1|t-1}F' + GQ^jG'
\end{aligned}$$

where

$$\begin{aligned}
x^{ij}_{t|t-1} &= E\left[x_t|Y_{t-1},s_{t-1}=i,s_t=j\right]\\
x^{i}_{t-1|t-1} &= E\left[x_{t-1}|Y_{t-1},s_{t-1}=i\right]\\
P^{ij}_{t|t-1} &= E\left[\left(x_t-x_{t|t-1}\right)\left(x_t-x_{t|t-1}\right)'|Y_{t-1},s_{t-1}=i,s_t=j\right]\\
P^{i}_{t-1|t-1} &= E\left[\left(x_{t-1}-x_{t-1|t-1}\right)\left(x_{t-1}-x_{t-1|t-1}\right)'|Y_{t-1},s_{t-1}=i\right]
\end{aligned}$$

for $i,j=1,2,3,4$. The problem that arises with four possible states is that the number of estimates for the state vector and their mean square error matrices grows exponentially with time. Indeed, at a given time t, we have t^4 estimates of the state vector to compute. The solution we adopt is to use the re-collapsing procedure suggested by Harrison and Steven (1976) which effectively provides re-approximations at each time t. These are given by:

$$x_{t|t}^{j} = \frac{\sum_{i=1}^{4} \Pr\left(s_{t-1}=i, s_t=j|Y_t\right) x_{t|t}^{ij}}{\Pr\left(s_t=j|Y_t\right)}$$

$$P_{t|t}^{j} = \frac{\sum_{i=1}^{4} \Pr\left(s_{t-1}=i, s_t=j|Y_t\right) \left\{P_{t|t}^{ij} + \left(x_{t|t}^{j} - x_{t|t}^{ij}\right)\left(x_{t|t}^{j} - x_{t|t}^{ij}\right)'\right\}}{\Pr\left(s_t=j|Y_t\right)}$$

The filtered estimate of the state vector is then obtained as:

$$x_{t|t} = \sum_{i=1}^{4}\sum_{j=1}^{4} \Pr\left(s_{t-1}=i, s_t=j|Y_t\right) x_{t|t}^{ij}.$$

A-3.1 Initial Values

Since one component of the state vector is non-stationary, we cannot initialize all components of the state vector and its covariance matrix to their unconditional expected values. The initial value we used are:

$$x_{0|0} = [0.95, 0, 0, 0, 0]'$$

and

$$P_{0|0} = \begin{bmatrix} 1e-5 & 0 \\ 0 & P \end{bmatrix}$$

where the submatrix P is given by

$$vec\left(P\right) = \left[I_4 - F1 \otimes F1\right]^{-1} vec\left(Q1\right)$$

with

$$F1 = \begin{bmatrix} \phi_1 & \phi_2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$Q1 = \begin{bmatrix} \alpha_1\sigma_{\xi 1}^2 + (1-\alpha)\sigma_{\xi 2}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\omega^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The initial value of the slope component β_t is set to 0.95, the average value of the slope for the pre-1973:1 period. We set its variance to a small number to reflect our prior that the trend function was stable before the break. The other components of the state vector are stationary and we use their steady state values as initial conditions. We experimented with different initial values for the first component of the state vector. The results were basically the same, except for the filtered trend function at the very beginning of the sample. The smoothed trend and cycle remain unchanged.

A-3.2 Constraints

A practical difficulty in the estimation of such Gaussian mixture models is the so called "label-switching problem" (see, e.g., Hamilton, Waggoner and Zha, 2002). This problem is due to the fact that the likelihood function $p\left(y_t|Y_{t-1}\right)$ does not change if the individual components of $p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right)$ are interchanged, and likewise for $p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right)$, so that

$$\begin{aligned} & p(y_t|s_{t-1},s_{t,}Y_{t-1})\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right)+p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i^*,s_t=j|Y_{t-1}\right) \\ = \; & p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i^*,s_t=j|Y_{t-1}\right)+p\left(y_t|s_{t-1},s_{t,}Y_{t-1}\right)\Pr\left(s_{t-1}=i,s_t=j|Y_{t-1}\right) \end{aligned}$$

Hence, we cannot identify the sates i and i^* without some normalization. To overcome this problem, we impose the restrictions discussed in the text.

A-3.3 Initial conditions and computations

To maximize the chances of obtaining parameter estimates that correspond to the global maximum of the likelihood function, we re-estimate the model 200 times with different initial values for the parameters that are drawn from a $N(0,3)$. The convergence criterion is set at $1e{-}4$ in the MATLAB command '`fminunc`'. We also repeated the same estimation procedure using a GAUSS code, setting the convergence criterion of the tolerance for a gradient change to less than $1e-5$. Since GAUSS trials gave better results, we used its output as the initial values for the MATLAB program, which then gave back the same answers.

Since this estimation and the filtering procedure are similar to the ones for Markov switching models, the basis for the construction of our computer codes was the GAUSS program written by Chang-Jin Kim (KIM_JE1.OPT) as discussed in Kim and Nelson (1999). The code is available from the book's website.

A-3.4 Smoothing

The smoothing algorithm used follows Kim (1994). The basic updating equations are:

$$x_{t|T}^{ij} = x_{t|t}^{i} + \widetilde{P}_t^{ij}\left(x_{t+1|T}^{j} - x_{t+1|t}^{ij}\right) \tag{A.4}$$

$$P_{t|T}^{ij} = P_{t|t}^{i} + \widetilde{P}_t^{ij}\left(P_{t+1|T}^{j} - P_{t+1|t}^{ij}\right)\widetilde{P}_t^{ij\prime} \tag{A.5}$$

where

$$\begin{aligned} x_{t|T}^{ij} &= E\left[x_t|Y_T, s_t = i, s_{t+1} = j\right] \\ P_{t|T}^{ij} &= E\left[\left(x_t - x_{t|T}\right)\left(x_t - x_{t|T}\right)'|Y_T, s_t = i, s_{t+1} = j\right] \\ \widetilde{P}_t^{ij} &= P_{t|t}^{i}F'[P_{t+1|t}^{ij}]^{-1}. \end{aligned}$$

The smoothed probabilities are given by:

$$\begin{aligned} \Pr\left(s_t = i, s_{t+1} = j|Y_T\right) &= \Pr\left(s_{t+1} = j|Y_T\right)\Pr\left(s_t = i|s_{t+1} = j, Y_T\right) \\ &= \Pr\left(s_{t+1} = j|Y_T\right)\Pr\left(s_t = i|s_{t+1} = j, Y_t\right) \\ &= \frac{\Pr\left(s_{t+1} = j|Y_T\right)\Pr\left(s_t = i|Y_t\right)\Pr\left(s_{t+1} = j|s_t = i\right)}{\Pr\left(s_{t+1} = j|Y_t\right)} \\ &= \Pr\left(s_{t+1} = j|Y_T\right)\Pr\left(s_t = i|Y_t\right) \end{aligned}$$

and

$$\Pr\left(s_t = i|Y_T\right) = \sum_{j=1}^{4}\Pr\left(s_t = i, s_{t+1} = j|Y_T\right).$$

As for the filtering procedure, we also use a re-collapsing approximation to avoid the problems given by the exponential increase in the number of states:

$$\begin{aligned} x_{t|T}^{i} &= \frac{\sum_j \Pr\left(s_t = i, s_{t+1} = j|Y_T\right)x_{t|T}^{ij}}{\Pr\left(s_t = i|Y_T\right)} \\ P_{t|T}^{i} &= \frac{\sum_j \Pr\left(s_t = i, s_{t+1} = j|Y_T\right)\left\{P_{t|T}^{ij} + \left(x_{t|T}^{i} - x_{t|T}^{ij}\right)\left(x_{t|T}^{i} - x_{t|T}^{ij}\right)'\right\}}{\Pr\left(s_t = i|Y_T\right)}. \end{aligned}$$

A-3.5 Numerical Integration for the slope of the trend function

To detect the break in the value of the slope of the trend function more accurately, we used the numerical integration method for non Gaussian (Gaussian Mixture) smoothing with numerical updating and filtering. These are basically the equivalent of the general formulae given by:

$$p(x_t|Y_{t-1}) = \int p(x_t|x_{t-1})\, p(x_{t-1}|Y_{t-1})\, dx_{t-1} \tag{A.6}$$

$$p(x_t|Y_t) = \frac{p(y_t|x_t)\, p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \tag{A.7}$$

where

$$p(y_t|Y_{t-1}) = \int p(y_t|x_t)\, p(x_t|Y_{t-1})\, dx_t$$

and

$$p(x_t|Y_T) = p(x_t|Y_t) \int \frac{p(x_{t+1}|Y_T)\, p(x_{t+1}|x_t)}{p(x_{t+1}|Y_t)} dx_{t+1}. \tag{A.8}$$

Equation (A.6) is the one step ahead prediction, or updating, (A.7) gives the filtered process, and the smoothing process is given by (A.8). For more details, see Kitagawa (1989). Our input is the filtered trend, $\{\widetilde{\tau}_t = \tau_{t|t}\}$, obtained using the algorithm described at the beginning of this section. We then estimate model (8). The parameters to be estimated are $\alpha_1, \sigma_{\widetilde{\eta}}, \sigma_{\gamma 1}$, and $\sigma_{\gamma 2}$. The initial values used are the same as before (see Section A3.1). After obtaining the parameter estimates, we compute (A.6), (A.7), and (A.8). The FORTRAN 77 programing code for these steps (except those for the parameter estimation) is provided by Genshiro Kitagawa from the Institute of Statistical Mathematics (see also Kitagawa, 1993). The Fortran code was translated to MATLAB code.

Table 1: Maximum Likelihood Estimates when a Change in Slope is Allowed

UCUR73			UC073			$ARIMA73(2,1,2)$		
	Estimate	s.e.		Estimate	s.e.		Estimate	s.e.
ϕ_1	1.328	(0.024)	ϕ_1	1.279	(0.045)	ϕ_1	1.522	(0.117)
ϕ_2	−0.418	(0.014)	ϕ_2	−0.373	(0.042)	ϕ_2	−0.601	(0.109)
μ	0.952	(0.026)	μ	0.951	(0.024)	μ	0.951	(0.021)
d	−0.288	(0.046)	d	−0.288	(0.043)	d	−0.287	(0.038)
σ_η	0.104	(0.007)	σ_η	0.000	(0.207)	θ_1	−1.283	(0.138)
σ_ε	0.843	(0.008)	σ_ε	0.945	(0.005)	θ_2	0.283	(0.137)
$\sigma_{\eta\varepsilon}$	0.088	(0.006)	$\sigma_{\eta\varepsilon}$			σ_e	0.936	(0.047)
$ln(L) = -280.505$			$ln(L) = -280.697$			$ln(L) = -278.930$		

Note: $ln(L)$ denotes the value of the maximized likelihood function.

Table 2: Maximum Likelihood Estimates of the Alternative Specifications

a) $ARIMA73(2,1,1)$			b) $AR73(2)$ in level		
	Estimate	s.e.		Estimate	s.e.
ϕ_1	1.279	(0.064)	ϕ_1	1.275	(0.064)
ϕ_2	−0.373	(0.064)	ϕ_2	−0.375	(0.064)
μ	0.951	(0.024)	μ	0.951	(0.023)
d	−0.288	(0.043)	d	−0.287	(0.041)
θ	−1.000	(0.013)	c	724.175	(1.592)
σ_e	0.945	(0.047)	σ_e	0.942	(0.047)
$ln(L) = -280.697$			$ln(L) = -281.201$		

Table 3: Simulation Results: Base Case.

$UCUR$				$UC0$				$ARIMA(2,1,2)$			
	Sample	Median	s.e.		Sample	Median	s.e.		Sample	Median	s.e.
ϕ_1	1.34	1.21	(0.29)	ϕ_1	1.53	1.44	(0.10)	ϕ_1	1.34	1.23	(0.25)
ϕ_2	-0.71	-0.52	(0.18)	ϕ_2	-0.61	-0.57	(0.10)	ϕ_2	-0.71	-0.44	(0.20)
μ	0.82	0.81	(0.02)	μ	0.81	0.80	(0.02)	μ	0.82	0.80	(0.02)
σ_η	1.24	0.97	(0.20)	σ_η	0.69	0.65	(0.18)	θ_1	-1.05	-0.88	(0.28)
σ_ε	0.75	0.94	(0.34)	σ_ε	0.62	0.65	(0.13)	θ_2	0.52	0.07	(0.25)
$\sigma_{\eta\varepsilon}$	-0.84	-0.59	(0.47)					σ_e	0.97	0.96	(0.05)
$\rho_{\eta\varepsilon}$	-0.91	-0.65									

Notes: 1) The values under the columns "sample" are reproduced from Tables 1-3, Morley et al. (2003); 2) "Median" and "s.e." denote the Median value and the Monte Carlo standard errors from 200 replications.

Table 4: Simulation Results: Effect of varying the change in slope d.

a) UCUR

	$d=-0.1$		$d=-0.4$		$d=-0.6$	
	Median	*s.e.*	Median	*s.e.*	Median	*s.e.*
ϕ_1	1.24	(0.26)	1.12	(0.28)	0.99	(0.29)
ϕ_2	-0.51	(0.16)	-0.48	(0.16)	-0.40	(0.18)
μ	0.90	(0.02)	0.75	(0.02)	0.65	(0.02)
σ_η	0.74	(0.37)	1.15	(0.16)	1.39	(0.15)
σ_ε	0.83	(0.28)	1.08	(0.33)	1.27	(0.38)
$\sigma_{\eta\varepsilon}$	-0.28	(0.43)	-0.95	(0.50)	-1.52	(0.65)
$\rho_{\eta\varepsilon}$	-0.46		-0.76		-0.86	

b) UC0

	$d=-0.1$		$d=-0.4$		$d=-0.6$	
	Median	*s.e.*	Median	*s.e.*	Median	*s.e.*
ϕ_1	1.32	(0.10)	1.47	(0.11)	1.51	(0.11)
ϕ_2	-0.45	(0.11)	-0.62	(0.11)	-0.67	(0.13)
μ	0.90	(0.01)	0.75	(0.02)	0.65	(0.02)
σ_η	0.31	(0.26)	0.75	(0.16)	0.84	(0.25)
σ_ε	0.86	(0.13)	0.55	(0.15)	0.48	(0.21)

c) ARIMA(2,1,2)

	$d=-0.1$		$d=-0.4$		$d=-0.6$	
	Median	*s.e.*	Median	*s.e.*	Median	*s.e.*
ϕ_1	1.26	(0.27)	1.15	(0.36)	1.03	(0.47)
ϕ_2	-0.45	(0.20)	-0.45	(0.23)	-0.33	(0.25)
μ	0.90	(0.01)	0.75	(0.02)	0.65	(0.19)
θ_1	-1.00	(0.29)	-0.80	(0.38)	-0.64	(0.49)
θ_2	0.09	(0.25)	0.13	(0.26)	0.12	(0.28)
σ_e	0.95	(0.05)	0.97	(0.05)	0.99	(0.05)

Table 5: Simulation Results: Filtered estimates of the residuals of the trend function $\eta_{t|t}$.

a) Sample mean and simulated values with $d = -0.29$

	Sample mean		Simulated	
	UCUR	UC0	UCUR	UC0
Full Sample	−0.003	0.004	0.003	0.012
Pre-break	0.159	0.148	0.144	0.136
Post-break	−0.168	−0.142	−0.139	−0.109

b) Simulated values for different values of d

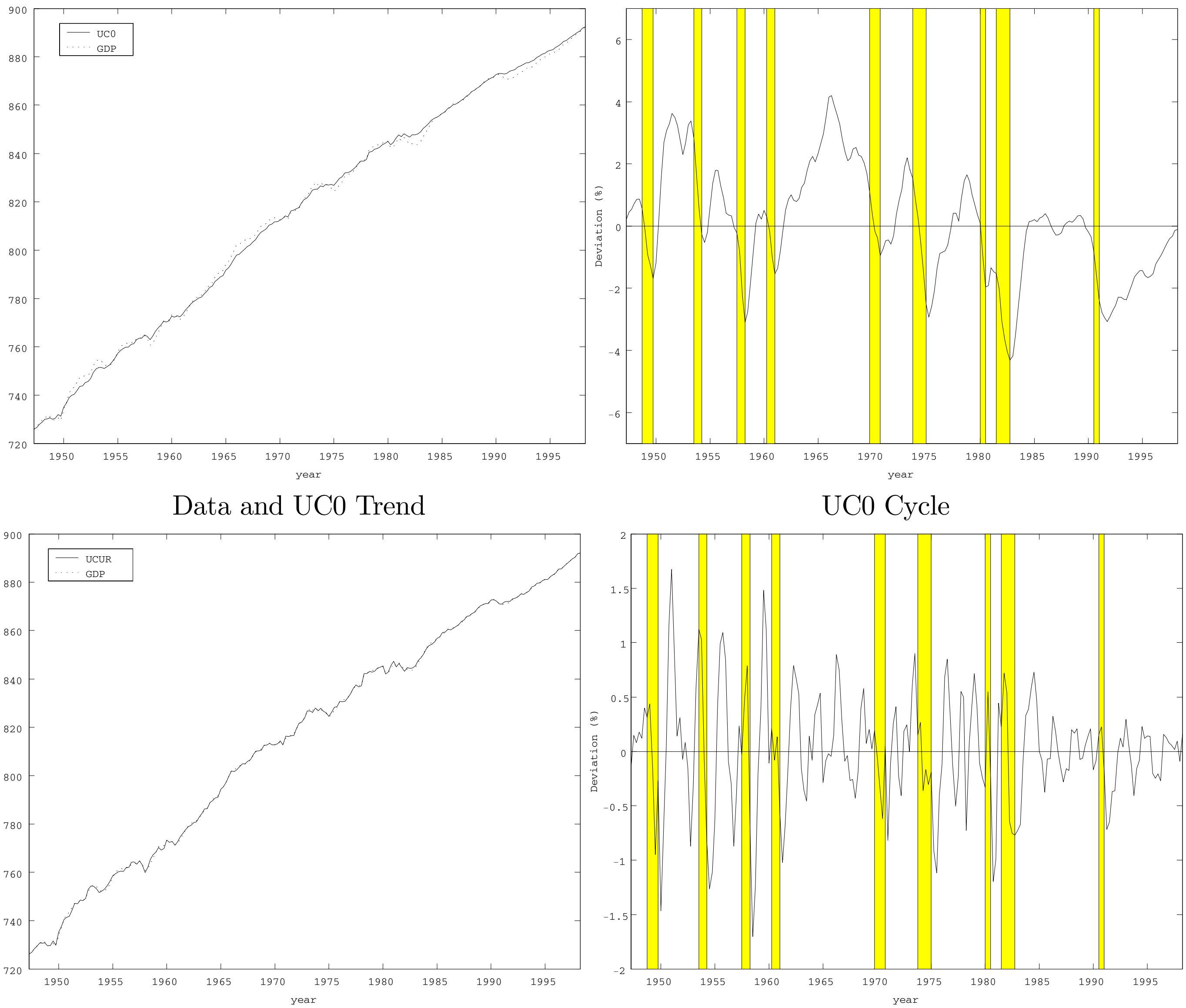
d	−0.1		−0.4		−0.6	
	UCUR	UC0	UCUR	UC0	UCUR	UC0
Full Sample	0.008	0.011	−0.002	0.004	−0.004	0.008
Pre-break	0.048	0.030	0.208	0.189	0.294	0.278
Post-break	−0.030	−0.009	−0.205	−0.176	−0.293	−0.254

Table 6: Parameter Estimates of the Gaussian Mixture Model

	Estimate
σ_η	0.0000
$\sigma_{\xi 1}$	0.2599
$\sigma_{\xi 2}$	1.2809
$\sigma_{\gamma 1}$	0.0001
$\sigma_{\gamma 2}$	0.0633
σ_ω	0.2495
ϕ_1	1.3795
ϕ_2	-0.4714
α_1	0.9000
α_2	0.5833
$\ln(L) = -272.4807$	

Table 7: Estimates for the Gaussian Mixture Model of the Filtered Trend

	Estimate
σ_η	0.4650
$\sigma_{\gamma 1}$	0.0000
$\sigma_{\gamma 2}$	1.0686
α	0.9992
$\ln(L) = -138.1379$	



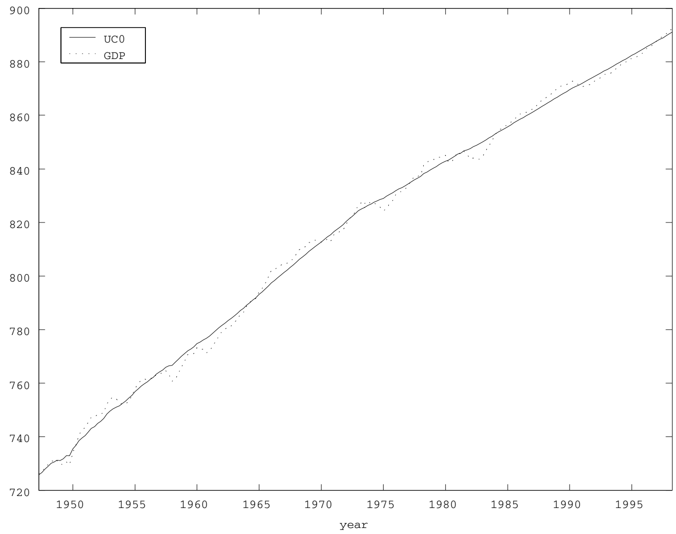
Data and UC0 Trend

UC0 Cycle

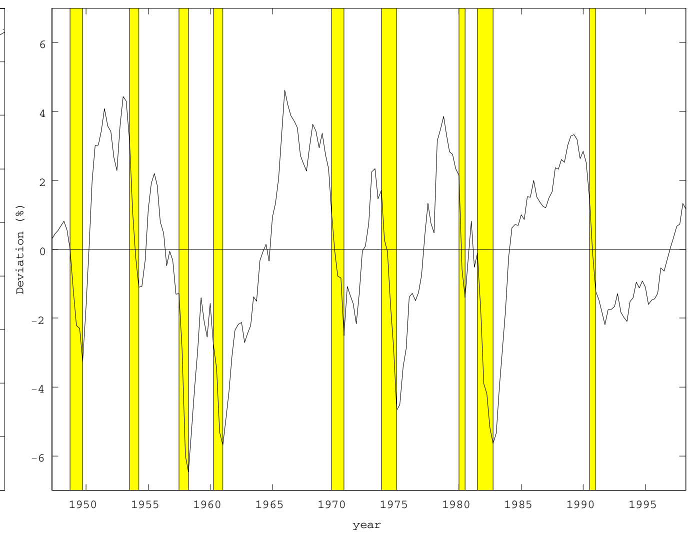
Data and UCUR-BN Trend

UCUR-BN Cycle

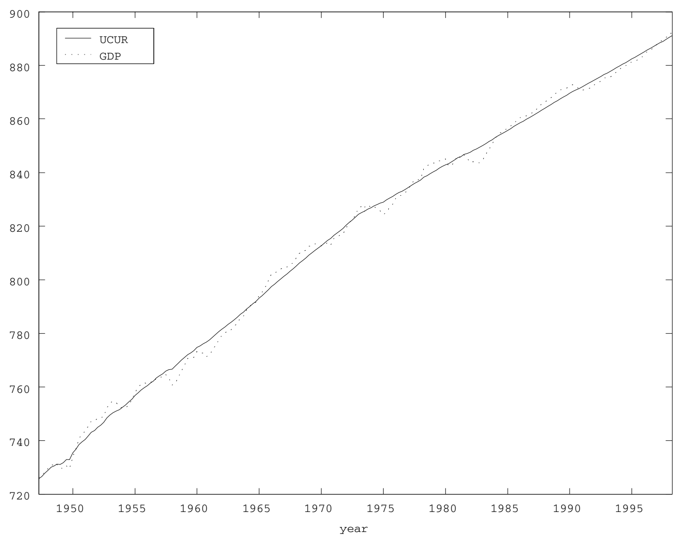
Figure 1: Trend and Cycle Decompositions of US log real GDP, 1947:1 -1998:2



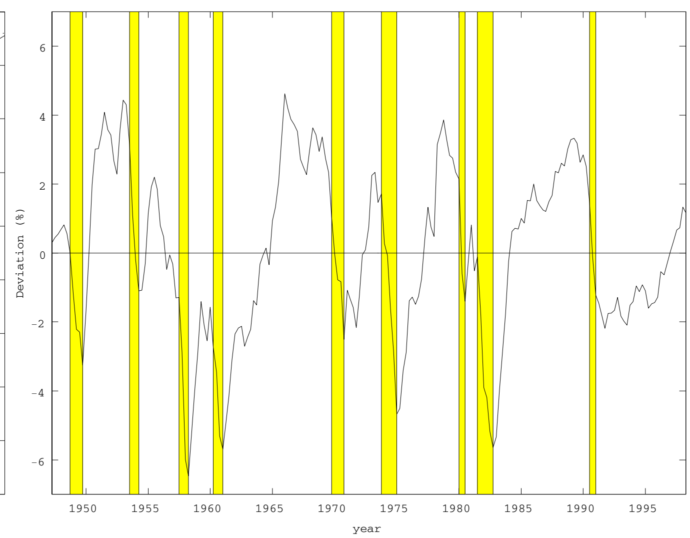
Data and UC073 Trend



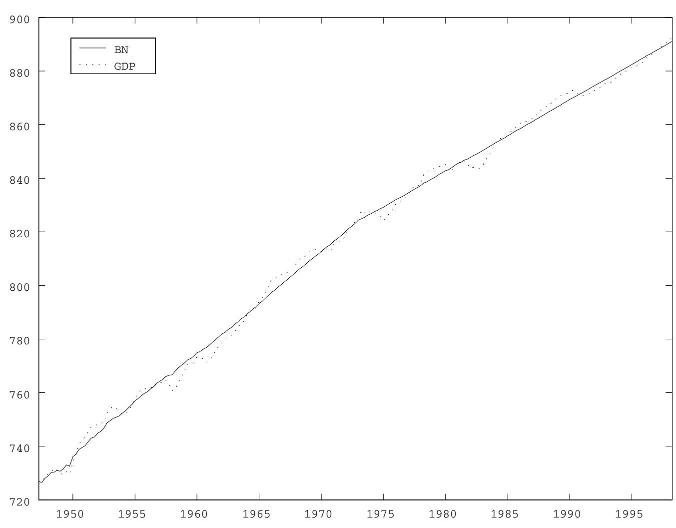
UC073 Cycle



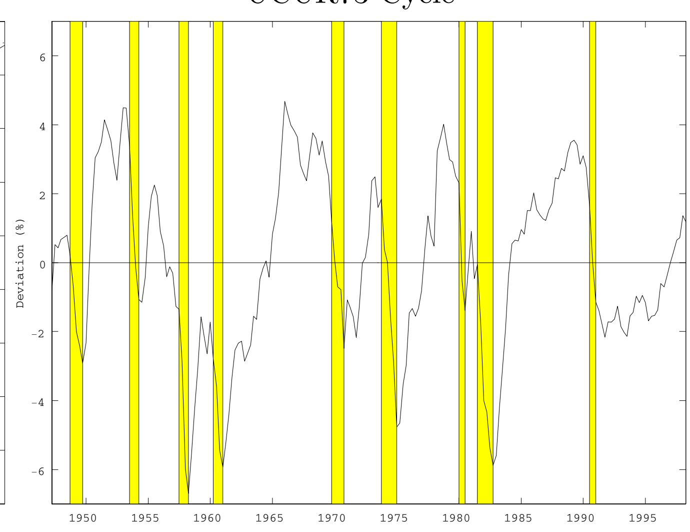
Data and UCUR73 Trend



UCUR73 Cycle

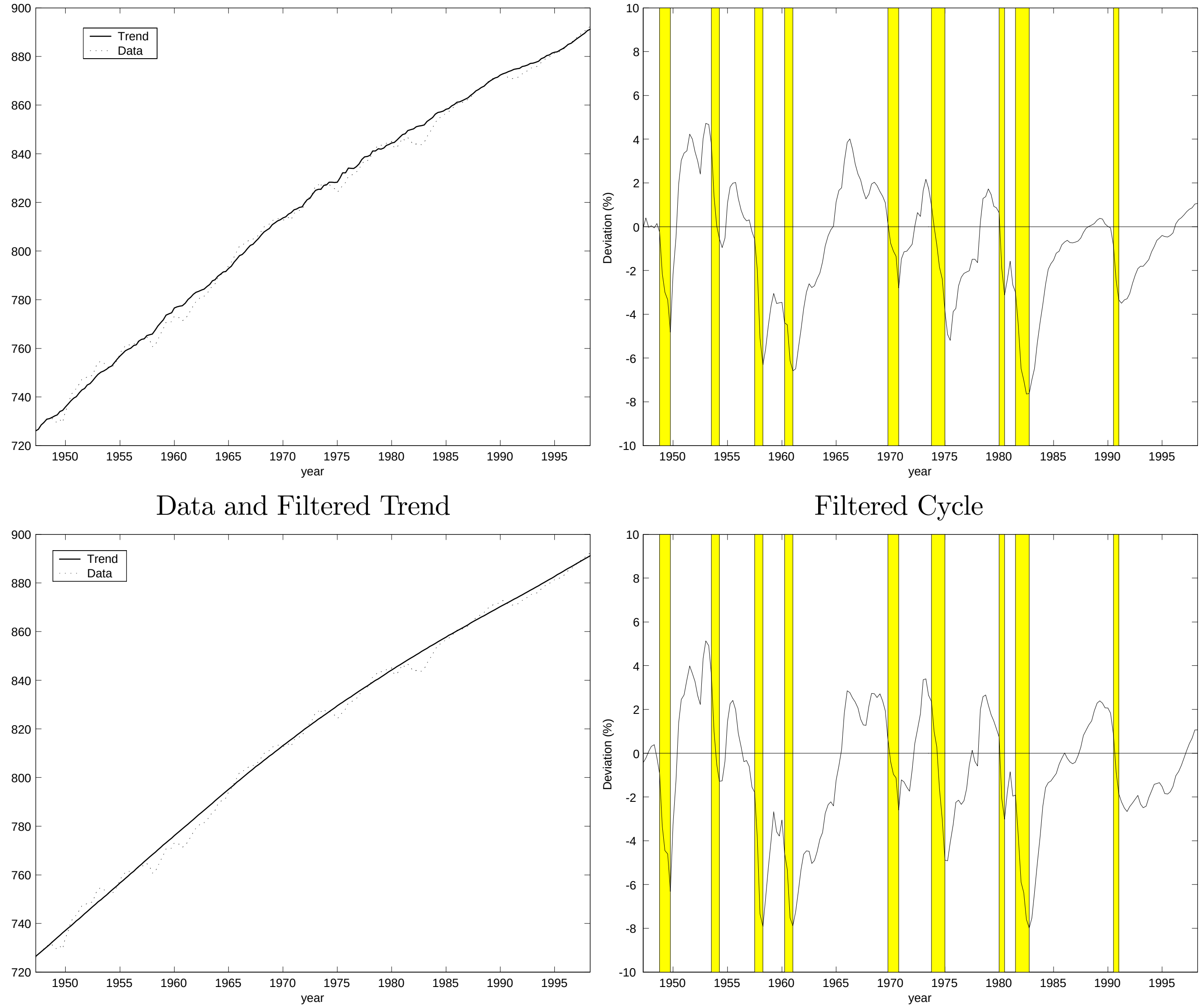


Data nd BN73 Trend



BN73 Cycle

Figure 2: Trend and cycle decompositions allowing for a change in slope in 1973:1



Data and Filtered Trend

Filtered Cycle

Smoothed Trend

Smoothed Cycle

Figure 3: Trend and cycle decompositions of the Gaussian mixture model

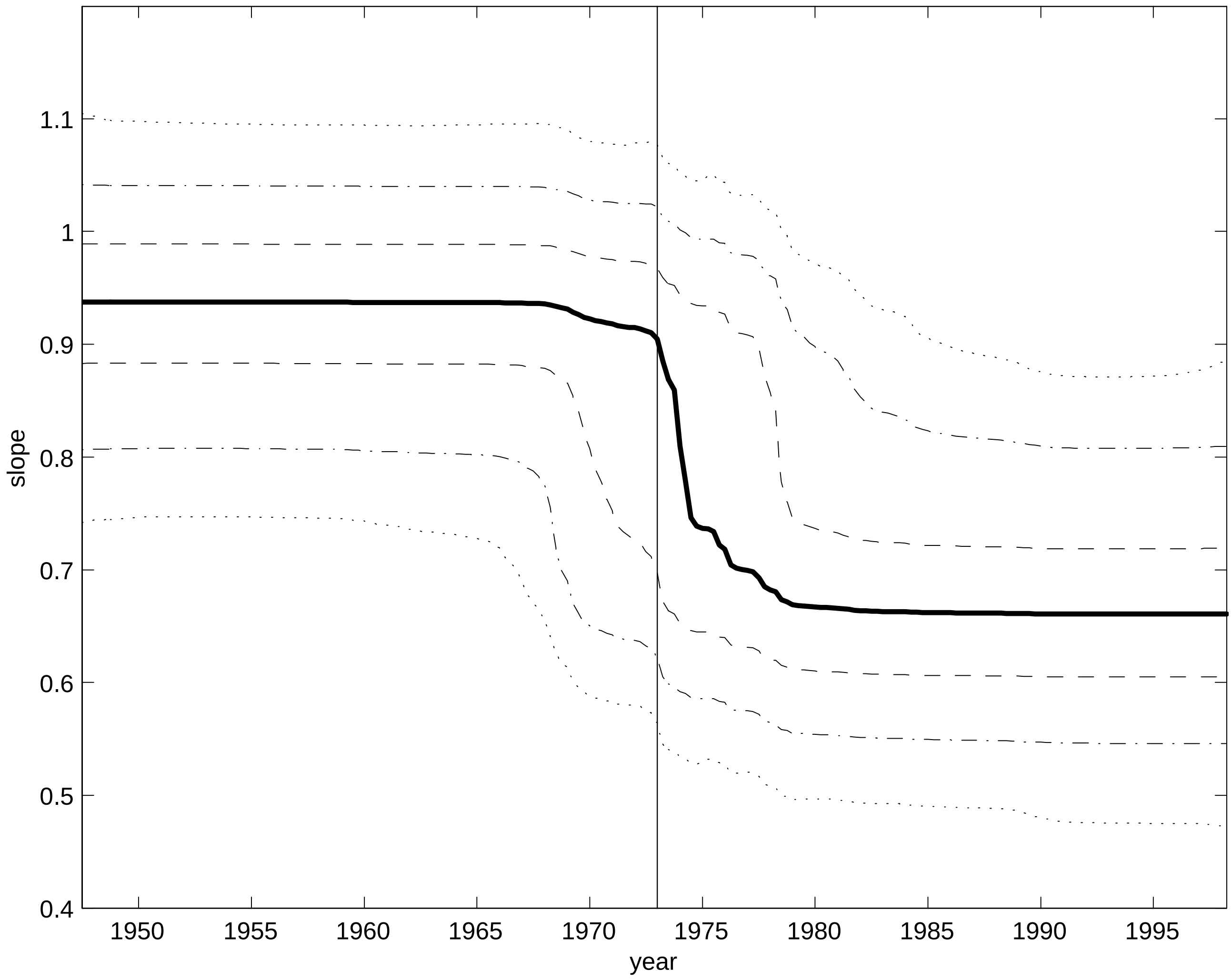


Figure 4: Smoothed estimate of the slope of the trend function

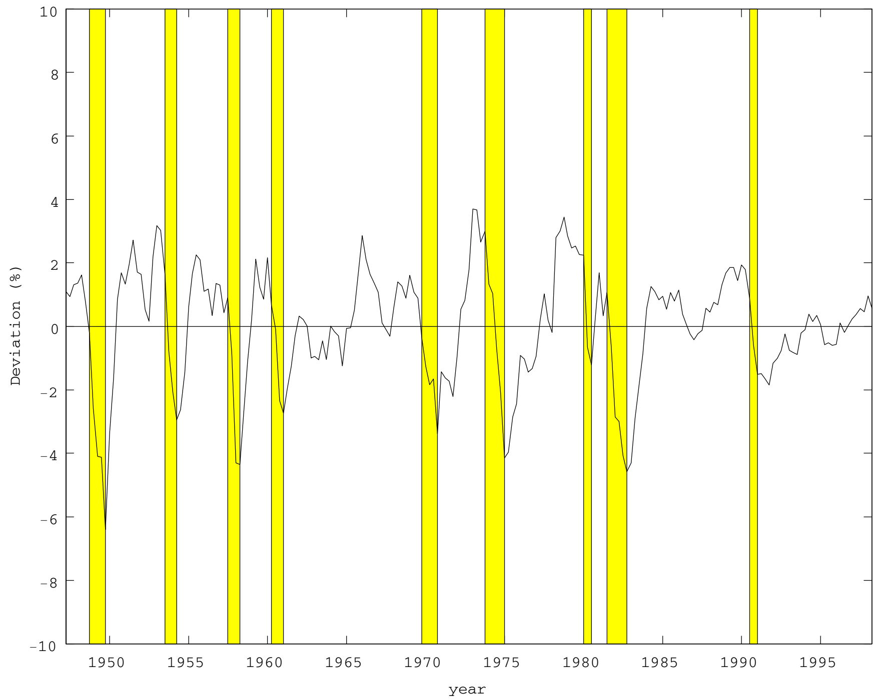


Figure 5: Hodrick-Prescott Cycle with $\lambda = 1,600$.

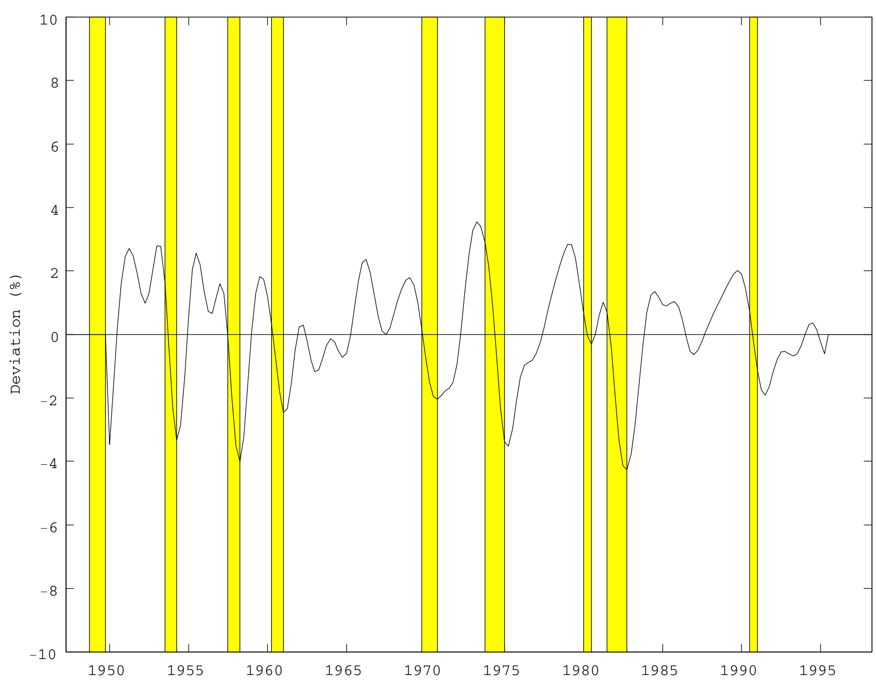
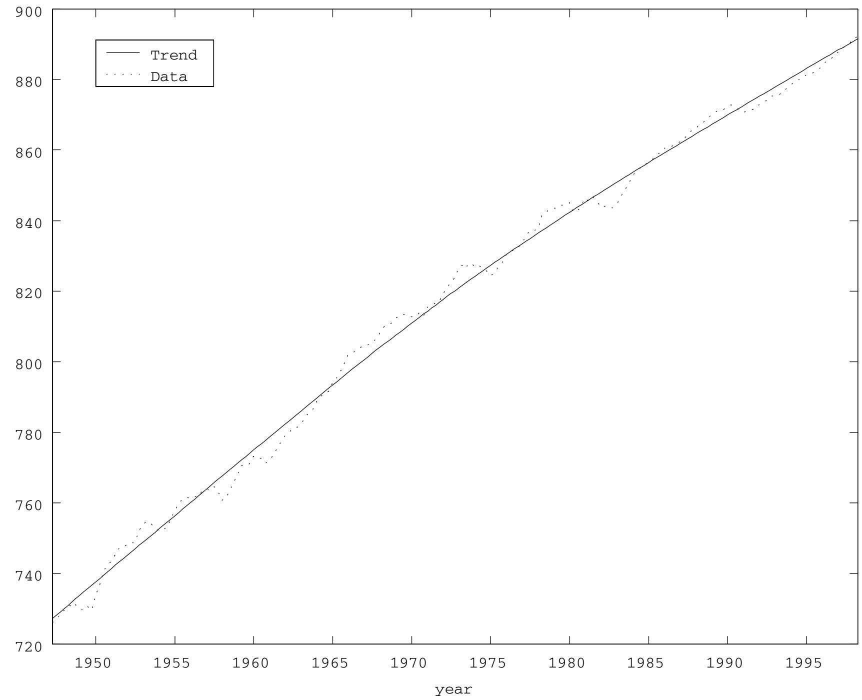
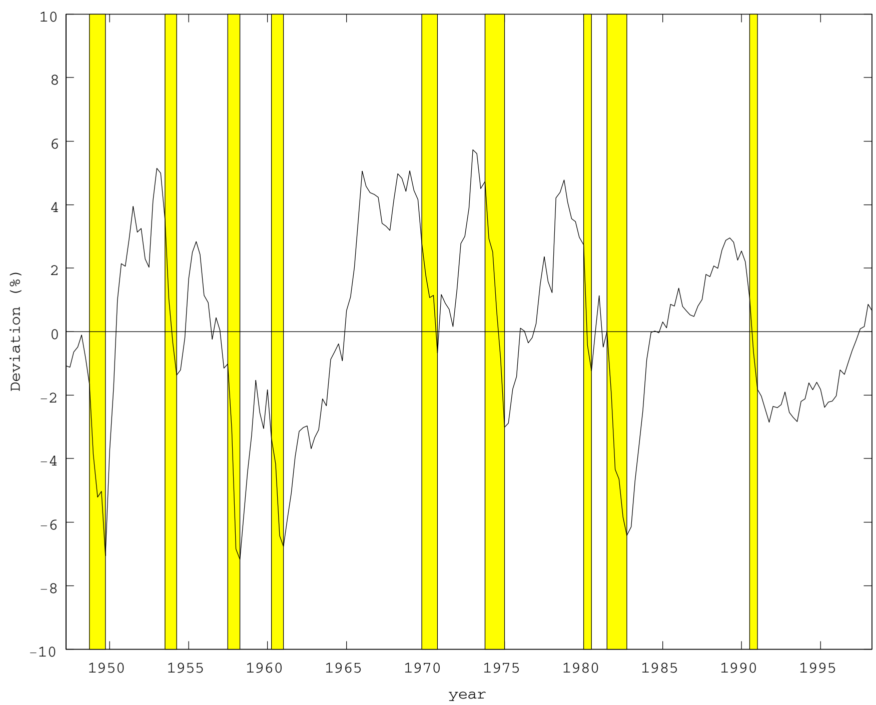


Figure 6: Cycle from the Band-Pass filter (6-32 quarters).



Data and Trend



Cycle

Figure 7: Trend-cycle decomposition from the Hodrick-Prescott Filter with $\lambda = 800,000$.