NBER WORKING PAPER SERIES

A HEURISTIC METHOD FOR EXTRACTING SMOOTH TRENDS FROM ECONOMIC TIME SERIES

Julio J. Rotemberg

Working Paper 7439
http://www.nber.org/papers/w7439

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1999

I wish to thank Michael Woodford for many thought-provoking conversations, John Driscoll for comments and the Harvard Business School Division of Research for support. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 1999 by Julio J. Rotemberg. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Heuristic Method for Extracting Smooth Trends from
Economic Time Series
Julio J. Rotemberg
NBER Working Paper No. 7439
December 1999
JEL No. E3, C1

ABSTRACT

This paper proposes a method for separating economic time series into a smooth component whose mean varies over time (the “trend") and a stationary component (the “cycle"). The aim is to make the trends as smooth as possible while also producing cycles with plausible properties. While the main justification for the method is intuitive, the method does a good job of separating these two components in some artificial examples where the constructed series are indeed the sum of smooth (possibly stochastic) functions of time and a low order autoregressive process. When the true trends consist of low order polynomials, the proposed method obtains trends that are of similar accuracy than fitted polynomial trends. In other cases, the MSE of the proposed trends is much lower. Similarly, except in quite special cases, the MSE of the proposed trend is considerably smaller than that obtained by the HP filter. VARs that involve the cyclical variables constructed by this method yield accurate representations of the behavior of the underlying cycles of several variables. By contrast, VARs with the series in differences give poor descriptions of the effect of cyclical shocks, even though Dickey-Fuller tests do not reject the hypotheses that the artificial series have unit roots. I apply the method to some well known aggregate time series. The results suggest that real wages in the U.S. are strongly positively correlated with military purchases and that the reduction in the growth of trend GDP in the U.S. started well before 1973.

Julio J. Rotemberg
Harvard Business School
Soldiers Field,
Boston, MA 02163
and NBER
jrotemberg@hbs.edu

This paper proposes a heuristic method for separating economic time series into a trend which is as smooth as possible and a well-behaved cycle. It then shows that this method does a good job of extracting trends and cycles in some artificial examples where the constructed series are indeed the sum of smooth functions of time and a low order autoregressive process.[1] Finally, it applies the method to some well known aggregate time series and discusses some of the properties of the resulting trends and cycles.

Many methods have already been proposed for separating stationary cycles from trends. Moreover, as Canova (1998) shows, the second moments of "cyclical" series extracted by different methods are not always the same. Different detrending methods thus emphasize different aspects of the data and one might thus conclude that there is no particular reason to prefer one method over another. Rather, one might prefer to keep several methods in mind whenever one is either analyzing data or testing models. This raises the question of whether it is useful to seek yet another method to separate trends from cycles.

I do so because I see several benefits of describing the data in terms of trends that are as smooth as possible subject to the constraint that the cycles be reasonably behaved. Smooth trends are ones whose changes in the rate of change are small. This means that trends that are as smooth as possible are relatively easy to describe. In particular, they tend to involve relatively long periods where changes in trend growth are zero or, at least, are of the same sign. This means that the question of what leads to changes in trend growth can be reduced to asking about what occurred during some well-defined "eras". If, by contrast, trend growth changes often and dramatically, one must find numerous individual "shocks" which affect trend growth.

The second advantage of making the trend as smooth as possible is that this ensures that the constructed cycle reflects a relatively large fraction of the shocks that have a strong short-run impact on the rate of growth of a series. The result is that the constructed cycles are well

[1]The use of computational examples to validate heuristic methods is used frequently in Operations Research. See, for example Winston (1994, p. 526 and 1050) as well as many recent issues of *Operations Research*. While it would clearly be better to have methods for which optimality can be demonstrated analytically, the importance and difficulty of the problem may justify the use of heuristics until better methods are developed.

suited to study whether short run changes in one series are related in short-run changes in another. Suppose, by contrast, that one constructs cycles by subtracting a relatively variable trend from the original series. It may then be the case that the constructed cycles of two series appear quite unrelated even if the short run changes in one series are associated with the short run changes of another. This could happen if the simultaneous short run changes in the two series are embedded in the trends.

In spite of these advantages of very smooth trends in terms of data description, such trends would have little meaning if, in truth, long run changes in economic time series were associated with sharp short run changes in these series. Smooth trends simultaneously embody all the long run changes while having only small changes in their short term rate of growth. They thus describe little of any importance if all substantial long term changes are associated with sharp short run changes. This criticism is of some importance given that many researchers have suggested that economic time series are well-described by stochastic processes containing a unit root. In these processes, temporary shocks have large long run effects and, by the same token, large long run changes are associated with shocks whose short term effect is nonnegligible. The dangers of over-interpreting "cycles" that are obtained by fitting smooth trends to stochastic processes of this type have been noted on several occasions.[2]

There are two responses one can make to this line of argument. The first is that the idea that long run changes in a series are not associated with large short run changes has some *a priori* appeal. This raises the obvious question of whether, if time series are indeed sums of smooth trends and cycles of short lived variability, the proposed method yields estimated trends that are close to the true trends. What makes this question difficult to answer is that there exist a great many stochastic and deterministic processes for trends which are smooth in the sense that the changes in their growth rates are small. Unless one knows the true trend process in advance, methods that are ideal for recovering one kind of

[2]For a recent example which focuses on the effects of applying the Hodrick-Prescott filter to series with unit rots, see Cogley and Nason (1995).

smooth trend while being quite poor for others would seem to be unattractive. Rather, a desirable method would obtain reasonable good estimates of the trend in a variety of different circumstances. I address the question of whether my method achieves this by considering a variety of numerical examples of smooth trends that are coupled to short-lived cycles.[3] I choose the examples so their behavior has some aspects in common with U.S. GDP.

Even when the true trend is close to being a low-order polynomial the accuracy of fitted polynomial time trends is similar to that of the trends obtained by my proposed method. In other cases, the proposed method is substantially more accurate than polynomial time trends. Similarly, it is more accurate than the HP filter except in circumstances that are quite special. This by no means proves the optimality of the proposed method. What it does show, and this is the main result of the paper, is that one can improve on existing methods.

A second response to the above criticism is that the fitting of unit root processes to series whose true trends are smooth can also lead to dramatically misleading inferences. Dickey-Fuller unit root tests fail to reject the hypothesis that a unit root is present in many of the artificial series I consider. At the same time, differencing creates the impression that innovations are associated with permanent changes even though the actual effect of shocks dissipates almost completely after 4 years. Even the short term impulse responses from a VAR estimated with differenced series provides a poor description of the actual effects of shocks on multiple series. Until improved unit root tests are developed, one is thus in danger of making misleading inferences whether one seeks smooth trends or fits unit root processes.

The idea that secular changes in economic time series ought to be represented by smooth lines is quite old. Bowley (1920, p. 137) says "The smoothed line now constructed [by freehand] represents the general tendency of the value of exports, when accidental and temporary variations are removed. If it were possible to separate entirely variations of short period from secular changes, to separate the ebb and flow of the tide of commerce from the steady current of increasing trade, we may suppose that we should obtain a result represented by this line. In it there are no sudden changes even in rates of growth... The direction

[3] For an earlier evaluation of detrending methods by simulating time series, see Kozicki (1999).

of the smooth line at any date may be called the *trend* of the series at that date" (italics in original).[4]

Several reasons can be given for the idea that long run changes in the value of a series (which must be part of the trend since they are not part of the temporary cycle) have only trivial reflections in the change in the rate of growth of the series at any particular point in time. Consider first the long term changes in the stock of an input such as labor. This is affected by demographic changes as well as by changes in attitudes towards work. Even important demographic changes have only gradual effects on the stock of workers, however. This is reflected in the changes of the quarterly growth rate of the U.S. population over 16 years of age. This series has some spikes. However, these seem to correspond mostly to measurement error because they tend to be followed immediately by spikes in the opposite direction. The post-war baby boom is reflected in this series by a number of positive values in the early 1960's but each of these is quite small.[5]

Just as in the case of labor, the stock of capital is smooth in part because very little of the existing capital is replaced in any given time period. Thus, fluctuations in investment from one period to the next have only tiny effects on the rate of growth of the change in the capital stock itself (see McCallum and Nelson (1997)). Since changes in the long run stocks of available labor and capital do not involve large short run changes, the same would be true for output if the long run changes in the technology used to produce output were also reflected in small short run changes.

The view that technology "in-use" is a smooth function of time does run counter to what is assumed in real business cycle models; these assume that there exist sharp changes in technology from one period to the next. However, students of actual technical progress such as Mansfield (1968) have emphasized that after someone discovers a new and improved process of production, it diffuses slowly at first. The cause of this might be that people only

[4]I obtained this source from Nerlove *et al* (1979) who note that almost all this text is already present in the 1901 version of this book.

[5]The largest movement in this period that is not offset by an immediate subsequent movement in the opposite direction is in the second quarter of 1962 where the quarterly rate of the population growth rose by two tenths of a percentage point.

learn slowly about the quality of the improvement (as in the model of Ellison and Fudenberg (1993)) or that they have a variety of attachments to old ways of doing things. In any event, if the adoption of technical improvements is gradual, the shift out in the production function due to technological progress is smooth as well. The result is that technical progress, which is the only source of long term productivity growth, may have only small effects on the rate of growth of productivity growth at any particular point in time.[6]

The notion that technological progress has only gradual effects naturally implies that cyclical movements (*i.e.*, ones who are not only temporary but also exhibit rapid changes) are due to other causes. This suggests that cyclical and trend changes ought to be orthogonal to one another. Consistent with this, my method requires that some aspects of the constructed trend's movements be orthogonal to some aspects of the resulting cycles. Strict orthogonality may well not be very important however, as I have constructed similar trends by replacing the orthogonality requirement by a requirement involving only the behavior of cycles. Once trends are smooth, a weakened version of my orthogonality requirement holds automatically since cyclical movements cannot take place at the same time as large changes in the trend if changes in trend growth develop only gradually.

The paper proceeds as follows. In Section 1, I discuss earlier approaches to extracting smooth trends from economic time series. These include polynomial time trends, the Hodrick-Prescott (HP) method and the use of band-pass filters. Section 2 then presents my method together with some intuition for why it can be expected to produce reasonable approximations in the case where the data generating process involves smooth trends. This proposed method is closely related to the Hodrick-Prescott method although it involves estimating the analogue of the parameter λ which they fix at 1600 for quarterly data. On the other hand, the method requires the input of two parameters. The first is the number of periods k such that, for a given degree of smoothness of the trend, the trend also minimizes the covariance of the cycle at t with the cycle at $t+k$. The second is the horizon v over

[6]On the other hand, long term changes in productivity could induce sharp short run changes in consumption and hours worked as a result of wealth effects. It remains an open question whether models of gradual technical diffusion can be constructed in which all macroeconomic variables evolve smoothly.

which one wants the estimated trend and cycle to be orthogonal.

Section 3 presents the artificial data which I study both with my method and with other approaches. These artificial data consist of the sum of smooth series and stationary autoregressive processes that return relatively quickly towards their mean. Section 4 evaluates the success of the method with the artificial data and compares it to the success of alternative methods which produce relatively smooth trends. Section 5 compares the inferences obtained from a VAR using artificial series detrended with the proposed method to the inferences that result if one's null hypothesis is that the series are difference-stationary.

Section 6 compares the results of detrending actual data with my proposed method to those of applying the HP filter. The two methods produce quite similar estimates of the second moments stressed in the real business cycle literature. On the other hand, the proposed method seems more revealing about the effects of changes in military purchases. In the United States, military buildups have often lasted several years so that these changes end up largely reflected in the HP trend. The result is that the correlation between detrended defense expenditures and output (or real wages) is weaker using the HP method than using my proposed method. Section 7 concludes.

1 Antecedents

Linear time trends are as smooth as possible since the rate of growth of these trends exhibits no changes whatsoever. However, linear detrending of economic time series is often problematic because the mean change in the log of many series differs across subsamples. This is notoriously true of U.S. GDP which exhibits faster growth in the 1947-1973 period than in the 1973-1997 period.

This has two consequences. First, it means that the trend value at t for many time series depends crucially on the sample over which the trend is estimated. Second, it means that when the whole series of observations is detrended with a single linear trend, the series remains either above or below the estimated trend for long periods of time. Thus, the estimated cycle (the difference between the series and the estimated trend) has very long swings. In other words, the estimated cycle is extremely persistent and appears close to having a unit root. This is illustrated for the log of U.S. GDP in Figure 1 which shows that this series has tended to be below its linear trend in the 1960's and 1980's while it has tended to be above this trend in the 1970's. This corresponds to no one's sense of booms and recessions since the 19760's are not generally regarded as a period of great prosperity nor are the 1960's remembered as periods where economic activity was seriously depressed.

There are several possible responses to these difficulties with linear trends, including simply abandoning the idea that economic time series can be described as the sum of smooth trends and relatively short lived cycles. One response along these lines is to suppose that trends are piecewise linear. The resulting discontinuities in the rate of growth of the trend mean that such trends are not particularly smooth overall, though they are extremely smooth within the periods where trend growth is assumed to be constant. What is unclear about these trends is what economic forces keep growth rates absolutely constant while allowing only for occasional, and dramatic, changes in growth rates. I thus focus on trends whose growth rates change gradually.

A traditional approach to constructing smooth trends is to model them as higher order polynomials of time.[7] This approach goes back at least to Moore (1919), who recommended that a cubic time polynomial be fitted by least squares. One problem with time polynomials is that it would require an unusual coincidence for their coefficients to remain stable over time. In part, this is because high powers of time become more important as the time index grows when additional observations are added to the sample.[8] In practice, the changes involved are quite dramatic. Figure 2 shows U.S. GDP as well as two estimated cubic trends. The first is estimated over the sample 1947:1 to 1998:1 while the second is estimated from 1947:1 to 1979:4, though the fitted value is extended over the whole sample. The line fitted over the shorter sample predicts trend declines in GDP at the end of the sample. In addition, the estimated trends in the middle of the 1970's differ by somewhat over 1.8%.

[7] See, for example, Malinvaud (1980, p. 455-456).

[8] In this respect using trends with fractional powers of time might be more desirable.

The second problem with polynomial trends is that one suspects that the timing of changes in the rate of growth of the trend is very sensitive to which powers of time one includes in the regression. Moreover, multicollinearity among different powers of time complicates the choice among these specifications.[9] This may be what led Bowley to aver that his freehand method was "more sensitive to changes in direction of the trend,"(Bowley, 1920, p. 137-8)

One solution to these problems is to use "moving average" methods which make the trend at t depend disproportionately on observations relatively near t. The well-known method proposed by Hodrick and Prescott (1997) falls in this category. It decomposes a time series whose observation at t is y_t (where t extends from 1 to T) into the sum of a trend series whose value at t is d_t and a cyclical series whose value at t is c_t by minimizing

$$\sum_{t=2}^{T-1}[(d_t - d_{t-1}) - (d_{t-1} - d_{t-2})]^2 + (1/\lambda)\sum_{t=2}^{T-1}(y_t - d_t)^2. \tag{1}$$

They recommend setting λ to 1600 in quarterly data. The first order conditions for this problem ensure that d_t is a symmetric moving average of y's with the highest weight being given to y_t. Since the approach minimizes the size of the changes in the rate of growth of the trend, the resulting trend is reasonably smooth. At the same time, the weight accorded in the objective function to the square of the cycle at t penalizes the existence of large cycles. This simultaneously ensures that the deviations from trend are relatively transitory and that d_t follows short term changes in y_t, at least to some extent. As King and Rebelo (1993) show, the influence of short term fluctuations of U.S. GDP on the HP trend is nontrivial. This can be seen in Figure 3 which shows the growth rate of the HP trend of U.S. GDP.

As one might expect, this shows that the trend growth rate has been lower after 1973 than before. However, the also shows that the 1980 and 1990 recessions have a profound effect on the HP trend. The result is that this trend is not nearly as smooth as a cubic trend fit to these data over the same interval. The mean value of the squared change in the HP's trend growth rate is 550e^{-10}. For the cubic time trend fitted over 1947:1 to 1998:1, the mean

[9] For example, quadratic trends change in the same direction from the beginning to the end of the sample, with the largest changes taking place at the edges of the sample.

value of square change in trend growth rate is about 70 time smaller; it equals $8e^{-10}$.

Baxter and King (1999) show that the HP filter is similar to a "band pass" filter which, assuming an infinite sample, lets through only oscillations whose period is longer than 32 quarters. Because economic time series contain only finite data points, one can only approximate pure band-pass filters. Nonetheless, Baxter and King (1999) show that the HP filter is almost identical to the best finite moving average approximation to this band-pass filter.

Christiano and Fitzgerald (1999) consider finite-sample ideal approximations to band-pass filters which do not involve constant coefficients. They compute optimal finite-sample filters based on IMA representations of the stochastic processes of economic time series and show that these are quite similar to the optimal filter in the case where the series can be represented by a random walk. This leads them to recommend the random walk-based band pass filter for economic applications.

The trends that result from the appealing requirement that they include all frequencies below that whose period equals 32 quarters are not as smooth as the HP trend. The mean square changes in the growth rate of the trend obtained by applying the Baxter-King (1999) method to U.S. GDP equals $1.2e^{-7}$, which is twice as large as in the case of the HP trend. This is surprising because the levels of the two trends track each other closely and even the second differences in the trends agree in their low frequency movements. The second difference of the Baxter-King trend has more high frequency changes, however. Using the Christiano-Fitzgerald (1999) approach the mean of the square changes in trend GDP growth is also equal to about 1.2^{-7}. While this too is twice as large as in the Hodrick-Prescott case, there is a difference between the two trends. Unlike what occurs in the HP and Baxter-King case, the difference between the Christiano-Fitzgerald trend's current value and the average of its value 5 quarters hence and 5 quarters ago is essentially uncorrelated with the Christiano-Fitzgerald cycle.

Given their frequency-domain interpretations, the lack of smoothness in these trends might be interpreted as suggesting that many shocks have large long-run consequences. In particular, the fact that only frequencies with periods longer than 32 quarters are included in these trends may suggest that shocks that have their effect almost exclusively within four years of their first appearance do not have much influence on these trends. If this were true, it would be reasonable to look for "noncyclical" shocks to explain the fluctuations in these constructed trends since business cycles are generally regarded as lasting only about four years.

However, it turns out that shocks whose effects are quite temporary can induce a great deal of variability at the frequencies that are included in these trends. To see this, consider a first order autoregressive process with first order serial correlation parameter equal to ρ. The fraction of this process' variance that accounted for by frequencies whose absolute value is below $z\pi$ equals

$$z + (2/\pi)\sum_{j\geq 1}\frac{\rho^j \sin(jz\pi)}{j}.$$

This means that, when ρ is equal to .75, 21% of the variance is accounted for by frequencies for which z is below 1/32. This is remarkable since, for this stochastic process, the effect of a shock after 16 quarters is about 1 percent of its original value.

A simple solution to this problem is to consider a band-pass filter whose trend contains only frequencies that are even lower than 1/32. It is difficult, however, to know *a priori* what the best cutoff frequency is. It would seem that in some cases, the linear trend is ideal and this can be thought of as a band-pass filter that lets through only the lowest possible frequencies. When this leads to implausible cycles, however, it makes more sense to allow the trend to contain somewhat higher frequencies. Thus, it is appealing to let the cutoff frequency depend on characteristics of the data. This requires a nonlinear filter in the sense that the weights on past and future observations must depend on characteristics of the data.

My conclusion from this brief review of previous approaches is that none give rise to reasonable cycles while having trends that are both smooth and stable with respect to the adding of new observations. This provides the motivation for the nonlinear filter I propose here. Before developing this approach, it is worth displaying the first difference of the trend I propose for U.S. GDP. I do this in Figure 4. This trend is quite smooth; the mean value of the square of the change in its growth rate equals $10.7e^{-10}$ which is only somewhat larger than the corresponding figure for the cubic trend estimated over the whole sample (which is $8.2e^{-10}$). Unlike this fitted cubic trend, however, it does not exhibit a monotone decline in trend growth.

As one would expect, my proposed trend grows more slowly after the early 1970's than before. Ignoring the observations near the boundaries of the sample which are less reliable, several things are interesting about this trend. The first is that the rate of growth of trend GDP actually rose in the early part of the sample. Second, the fastest reduction in trend growth took place in the early 1970's, particularly around 1973.[10] However, the figures also show that the rate of growth of the trend actually started falling considerably earlier. Indeed, the highest level of trend growth is seen in the last two quarters of 1962. This suggests that, if one accepts the notion that trends ought to be smooth, one might be led to interpret the post 1970 secular slowdown of the U.S. as having been due to events in the early 1960's, including possibly the policies associated with the Great Society.

2 The Proposed Method

Given the series y_t, I propose to construct its trend d_t by minimizing

$$\sum_{t=2}^{T-1}[(d_t - d_{t-1}) - (d_{t-1} - d_{t-2})]^2 + (1/\lambda)\sum_{t=1+k}^{T}(y_t - d_t)(y_{t-k} - d_{t-k}) \tag{2}$$

where the parameter λ is set at the lowest possible value that ensures that

$$\sum_{t=k+v}^{T-k-v}(y_t - d_t)[(d_{t+v} - d_t) - (d_t - d_{t-v})] = 0. \tag{3}$$

The procedure thus requires the input of two integer values, k and v. It is somewhat similar to the procedure advocated by Hodrick and Prescott (1997) since (1) is obtained from (2) by setting k to zero and λ to the value of 1600 for quarterly data. My procedure differs from theirs only in that I set k to a strictly positive value (generally 16) and that I choose λ through the independence requirement (3) - which still requires a choice of v.

[10]This fits with the findings of Perron (1989). If, like him, one imposes the *a priori* assumption that trend growth must be constant except for having some discontinuous jumps, the figure suggests that one ought to find that a jump occurred at around this time.

For a given value of λ, the first order condition for minimizing (2) when t is between $1+k$ and $T-k$ are

$$\begin{aligned}[(d_{t+2}-d_{t+1})-(d_{t+1}-d_t) &- 2[(d_{t+1}-d_t)-(d_t-d_{t-1})]+[(d_t-d_{t-1})-(d_{t-1}-d_{t-2})] \\ &= \frac{1}{2\lambda}\Big[(y_{t+k}-d_{t+k})+(y_{t-k}-d_{t-k})\Big] \qquad (4)\end{aligned}$$

To keep the analysis simple I use these first order conditions for t going from 1 to T by setting $(y_{t+\ell}-d_{t+\ell})$ equal to zero whenever $t+\ell$ exceeds T or is smaller than 1 because ℓ is negative. Similarly, I set the terms in square brackets $[(d_\ell-d_{\ell-1})-[(d_{\ell-1}-d_{\ell-2})]$ equal to zero whenever ℓ is either smaller than 2 or greater than T. I am thus supposing that, outside the sample I observe the expected rate of change of the trend does not change and the cycle is expected to be zero. While other initial conditions that use sample information might provide an improved fit, their exploration is left for future research.

The result is that the vector of trend estimates d is a linear function of the vector of observations y

$$Ad = By$$

where the matrices A and B depend only on λ and k. This implies that we can obtain the trend as

$$d = A^{-1}By. \qquad (5)$$

For given values of k and λ, the trend is thus obtained by a linear filter. This filter belongs to the HP family when k is zero. Even when k takes the value of 16 which I favor below, both filters are similar when λ is high. This is illustrated in Figure 5 which gives the approximate gain both filters would have if λ were set to equal to 70000 and the filtered series were stationary. The subtle but important differences between the filters that make a higher value of k preferable are not visible in the figure which shows mainly that the are extremely similar low-pass filters.

I now provide some reasons to suggest that this procedure has some a priori appeal. In effect, the procedure minimizes the expected square changes in the change of the trend subject to the condition that the covariance between the cycle at t and the cycle at $t-k$ be smaller than a pre specified quantity ν which is, in turn, chosen to ensure that the change in the trend over some horizon be orthogonal to the cycle at t.[11]

Both the pre-War and the post-War NBER chronologies suggest that business cycles are quite short. Between the NBER trough of 1949:4 and the trough of 1991:1, there were seven other troughs so that troughs have been separated by an average of 5 years. With one exception (in 1982) these troughs were discrete cyclical events that took place when the effect of the previous recessions was felt to have been completely dissipated. Similarly, the mean duration for the 21 pre-War United States cycles reported in Burns and Mitchell (1946 p. 78), where this duration includes both the contraction and the expansion, is only about four years. These considerations suggest that the cyclical component of aggregate time series at t ought to have a small, possibly even a negative, covariance with the cyclical component four years hence. This provides a motivation for setting k equal to 16. Even if this exact value of k is somewhat arbitrary, it seems clear that supposing that the business cycle is temporary implies that the covariance of the cycle at t with the cycle at $t+k$ ought to be small for a value of k which is similar to the length of time that cyclical shocks have their effect.

Suppose that, by contrast, k is set to a substantially smaller value. Then, the second term in (2) is reduced by having the trend track cyclical movements. If the trend did not do this, a positive cyclical shock at t would raise both $(y_t - d_t)$ and $(y_{t+k} - d_{t+k})$ and thereby penalize the objective function. The same logic explains why setting k to zero tends to lead the trend to track short term movements in the series.

The foregoing discussion explains why having a relatively small covariance between $(y_t - d_t)$ and $(y_{t+16} - d_{t+16})$ is desirable but does not rationalize (2). One role of this orthogonality condition is to ensure that the trend has a desirable degree of smoothness. To see this, suppose first that λ is set at a high value.

[11] I have also explored an alternate procedure which simply sets ν equal to zero when k equals 16. This has some appeal if, as I suggest below, one believes that cyclical shocks have negligible effects after 16 quarters. The results from this alternate procedure are fairly similar to those I obtained using (3), although the mean square error from this alternate procedure was higher for the artificial series I present in the next section.

If λ is essentially infinite, the d_t that solves (2) is a linear trend and (3) is satisfied because its second term is zero. For somewhat lower values of λ, the expression in (3) is negative unless the true trend is linear. The reason for this is the following. When the true trend does not have a constant slope and λ is high but not infinite, the constructed trend tracks the actual trend somewhat. By doing this tracking, the constructed trend reduces the size of the constructed cycles and this reduces the second term in (2) (Indeed, this is the reason that even in the limit where λ is arbitrarily large, the constructed trend is not an arbitrary straight line but a straight line that stays as close as possible to the actual points in the sample.) However, because λ is large so that the constructed trend is close to being a straight line, it still remains true that as in Figure 1 observations that are located where the slope of the true trend is falling find themselves classified as being cyclical peaks. Since the constructed trend's slope is also falling at these observations, positive values of the cycle are associated with reductions in the trend rate of growth so that the expression in (3) is negative.

Now suppose that λ is set at a low value. Then, the minimization of (2) is directed mainly at making the cycle at t negatively correlated with the cycle at $t-k$. The method achieves this by making the trend a quasi periodic function where peaks at t are matched with valleys at $t-k$. At points in time when the trend is relatively high, the cycle is measured to be negative and the future rate of growth of the trend is low relative to the past rate of growth of the trend (because of the quasi-periodic nature of the trend). This means that low values of λ ensure that the covariance between one's estimate of the cycle at t and the change in the growth rate of the trend is positive. The same is thus true for the sum in (3).

Between the region where λ is sufficiently low that the expression in (3) is positive and the region where λ is so high that it is negative, one expects to find a λ such that (3) is actually satisfied. At this point one has a trend that is smooth enough not to have unnecessary quasi-periodic movements about the true series but which, at the same time, is not so smooth that there are long periods where the cycle is either above or below the trend.

With k sufficiently high that the trend does not track the short term movements in the series, the resulting trend ought to be quite smooth. As a result, the expression in (3) ought to be close to zero for a variety of different values of v even if it is exactly zero only for one of them. Thus, the resulting trend ought to be relatively insensitive to the choice of v.

If k is small, by contrast, the minimization of (2) encourages the trend to follow the short term movements in the series. As discussed above, this tracking means that high values of the cycle are associated with temporarily high values of the trend so that the expression in (3) tends to be negative. Still, one can ensure that (3) holds by lowering λ and thereby inducing some quasi-periodic movements of the trend around the true series. However, in this case (3) is an artifact of combining quasi-periodic movements in the trend, which make the expression in (3) positive, with the tracking of short term movements by the trend which makes the expression negative. Both of these effects are presumably quite sensitive to v. This means that, with low k, one would not expect the value of λ which makes (3) hold for a given v to lead to low values of (3) for nearby choices of v.

In my numerical experiments with the series I discuss in the next section, this pattern emerged. With high values of k (such as my preferred value of 16), the constructed trend had the property that the expression in (3) was close to zero for a large range of v's. Moreover, the constructed trend was quite insensitive to v. Neither of these properties held when k was set to a small value.

With k set to zero, one is effectively choosing among HP trends with different values of λ. However, it is then impossible to obtain trends that minimize (2) while satisfying (3). Even when λ is quite large, (2) gives some weight to the variance of the cycle. The result is that the trend rise whenever the cycle is positive and the expression in (3) is negative.

The ability to obtain trends and cycles whose short term movements are unrelated by setting $k = 16$ should not obscure the fact that the construction of trends and cycles that are independent at all leads and lags is essentially impossible. To see this, suppose that the function that gives d_t is given by $f(y_{t-\ell})$ where this should be thought of as a function that depends on observations for many different values of ℓ. Note also that my method lets f be a constant linear function of this type except at the boundaries. Then, given that y_t

equals $c_t + d_t$ by construction, we must have $d_t = f(d_{t-\ell} - c_{t-\ell})$ so that there is a stochastic singularity connecting the leads and lags of c to the leads and lags of d. The fact that they both are constructed from the same set of observations makes it impossible for them to be truly independent.

To get a feeling for the extent to which (3) is violated for alternate values of v, I construct "regression coefficients" which give the amount by which the trend is expected to change for a given level of $-c$. In particular, I compute β_v which is given by

$$\beta_v = -\frac{\sum_{t=k+v}^{T-k-v}(y_t - d_t)[(d_{t+v} - d_t) - (d_t - d_{t-v})]}{\sum_{t=k+v}^{T-k-v}(y_t - d_t)^2}.$$

This gives the amount by which the trend increase between t and $t+v$ is supposed to differ from the trend increase from $t-v$ to t for a given level of $-c$. If the series is in logs so that $-c$ measures the percentage by which the series is below its trend value, this gives the percent by which the trend increase in the future is supposed to exceed the trend increase in the past. A value of one would then say that the expected increase in the trend over v quarters is equal in magnitude to the total expected increase in the cyclical component (since this always returns to its mean of zero).

I now turn to the description of the artificial data which I analyze with this method.

3 Artificial Data

I generate the observations of y_t by summing together two artificially generated time series D_t and C_t. In each case, I generate 205 observations so that the length T of the artificial data is the same as that of the quarterly U.S. GDP and hours data which I use below. The first cyclical series I consider, C_t^i is a first order autoregression

$$C_t^1 = .75C_{t-1}^1 + \epsilon_t^1$$

where ϵ_t^1 is a drawn from a standard normal distribution using a random number generator. The series C_t^2 is, instead, given by a second order autoregressive process

$$C_t^2 = 1.3C_{t-1}^2 - .4C_{t-2}^2 + \epsilon_t^2$$

where the ϵ_t^2's are separate pseudo-random draws which are generated in the same way as the ϵ_t^1's. This somewhat more persistent process is inspired to some extent by the behavior of U.S. GDP. When U.S. GDP from 1947:1 to 1998:1 is explained by two lagged values as well as a linear and a quadratic trend, the coefficients on the two lags of GDP are close to 1.3 and -.4. My baseline stochastic process for the cycle, C_t^3, is the sum of C_t^1 and C_t^2. The advantage of using a cycle that is the sum of distinct two random variables is that one is then able to see how the presence of trends distorts the estimated relationships between two variables. Except for the case of a trend that I choose specifically because it makes the HP filter optimal, I use just these three cyclical series.

My first trend is a simple linear function of time. In particular, I set

$$D_t^1 = .845t \qquad\qquad t = 1, \ldots, T.$$

The parameter in this linear trend is chosen so that the sample standard deviation of the trend is 50. I ensure that the trends D^1 through D^6 have this same standard deviation. This standard deviation is about 16 times the sample standard deviation of C_t^3 (while the sample standard deviations of C_t^1 and C_t^2 are somewhat smaller). Making the sample standard deviation of my trends be 16 times larger than that of the my baseline cyclical series is again motivated by the behavior of U.S. GDP. If one fits a quadratic trend to the log of U.S. GDP, the sample standard deviation of this estimated trend is about 16 times larger than the sample standard deviation of the residual.

My second trend is quadratic. It is given by

$$D_t^2 = 1.16t - .0015t^2.$$

I consider also three trends which involve trigonometric functions. I do this because these functions are quite smooth while also allowing for a quite rich array of nonlinear features. The first of these, D_t^3 is an extremely smooth monotone, convex and increasing function of time. It is given by

$$D_t^3 = 225\cos\left(\frac{t}{200} + 3.5\right).$$

While not exactly a polynomial of time, this can be well approximated by a cubic trend. In particular, a regression of this trend on time, time squared and time cubed has residuals whose standard error is only .03 (when, as I said above, the standard error of the left hand side is 50).

The next trend, D_t^4, is less similar to a cubic polynomial of time. It is given by

$$D_t^4 = .687t + 2.29(\sin(t/50) - \cos(t/50)).$$

This trend is depicted in Figure 6. The linear coefficient on t ensures that trend growth is always positive. Inspired by the trend shown in Figure 4, the rate of growth of this trend first rises and then falls. The trend growth rate of D^4 also rises somewhat at the end of the sample. The R^2 from regressing this trend on time, time squared and time cubed is just under .9992. Thus, while this is further than D^3 from being a cubic polynomial of time over this restricted sample, it is still fairly close to a polynomial of this sort.

My fifth choice of trend is meant to challenge my method by making the "trend" have "growth cycles". In particular, it is given by

$$D_t^5 = .87t + 17.4 \cos\left(\frac{t}{20}\right).$$

The second term in this expression has a period of about 31 years while the first term, again, ensures that the function itself is monotone. While the periodicity of the cycles in this trend is much higher, the mean of the square of the second difference in this trend is actually somewhat lower than that for D^4. It equals $9.9e^{-4}$ rather than $13e^{-4}$.

The five trends I have described so far are deterministic and this feature has two advantages. The first is that these functions are extremely smooth while the second is that they are easy to describe completely. However, my method of recovering trends would not be very interesting, if it could only be applied successfully to trend variations that are completely determined by initial conditions. The idea that trends are both smooth (so that their rate of change varies slowly) while their long run level is hard to determine from the samples we have available (so that long run forecasting is hazardous) suggests that, indeed, initial conditions may be important because "shocks" to the trend have their most important effects only after considerable time elapses. But this need not rule out the existence of some effect of the shocks that take place during the sample.

I thus also consider two stochastic trends. The first of these is extremely smooth. In particular, I let the trend D_t^6 satisfy

$$[(D_t^6 - D_{t-1}^6) - (D_{t-1}^6 - D_{t-2}^6)] - [(D_{t-1}^6 - D_{t-2}^6) - (D_{t-2}^6 - D_{t-3}^6)] = \epsilon_t^5$$

where ϵ_t^5 is drawn from a normal distribution whose standard deviation is $8.2e^{-5}$. The randomness of the third difference of the trend implies that shocks with trivial short term consequences have huge long run effects. To ensure that only the sample noise makes the trend complicated to estimate, I suppose that, before the sample starts, D^6 grows at the constant rate of .76. Thus, without the noise, this series would be a simple linear trend. Since I only look at one sample path of this trend, it seems worth displaying this in Figure 7. While the series appears to be a linear trend, its rate of growth has a sample standard deviation of .08. The mean in the second half of the sample is .92 while that in the first half is only .78. Still, there is a sense in which this trend is very smooth, since the mean of the square second difference in this trend equals $2e^{-6}$.

Finally, I also consider a stochastic trend whose second difference is an i.i.d. normal variable. In particular D_t^7 satisfies

$$(D_t^7 - D_{t-1}^7) - (D_{t-1}^7 - D_{t-2}^7) = \epsilon_t^6.$$

I constructed a single path of this trend by drawing ϵ^6 from a standard normal with unit variance. As it happens, this path is monotone increasing and its variance is quite large. This trend is also not very smooth. Even if I had multiplied the resulting trend by a constant to make the standard deviation of D^7 be equal to 50, the mean of the square second difference in this trend equals .002.

The main reason for analyzing this trend (which is not as smooth as the others) is that, as shown by Hodrick and Prescott (1997), the HP filter leads to optimal estimates of the trend when the true trend is given by D^7 as long as the cycle is given by an i.i.d. random variable whose standard deviation is forty times larger than the standard deviation of e_t^6. I thus couple this trend with a cycle C_t^4 which equals the independently identically distributed random variable e_t^4 whose standard deviation is 40.

I also couple the trend D^7 with two more persistent cycles C_t^5 and C_t^6. The former follows the same AR(1) process as C_t^1 but its residual equals .66 times e_t^4. This ensures that the standard deviation of this cycle is roughly the same as the standard deviation of C_t^4. Similarly, the cycle contained C_t^6 follows the same AR(2) process as C_t^2 but its residual equals .3 times e_t^4 so that the standard deviation of this cycle is also close to that of C_t^4.

I thus have seven possible trends and four different types of cycles (i.i.d., AR(1), AR(2) and sum of AR1) and AR(2)), though I do not consider all possible combinations. I denote series whose trend is D_t^i and whose cycle is C_t^j by y_t^{ij} so that

$$y_t^{ij} = D_t^i + C_t^j.$$

4 The accuracy of detrending methods applied to artificial data

I start by analyzing the extent to which various detrending methods, including the one I propose with $k = 16$ and $v = 5$, recover trends that are similar to the trends contained in the artificial data. Because the sum of the estimates of c_t and d_t equal y_t just as does the sum of C_t and D_t, the error in estimating the cycle, $c_t - C_t$, is equal to $D_t - d_t$. This means that the mean square error of the estimated trend is the same as the mean square error of the estimated cycle. Table 1 presents these mean square errors for various detrending methods and various artificial data series. I compute the MSE's (as well as all other statistics concerning these artificial series) ignoring the 16 first and last observations. My proposed method is often substantially less accurate at these boundaries. It may be possible to improve on this by using better initial conditions. However, the problem may well be inherent in the estimation of extremely smooth trends. When trends are very smooth, moving average methods only lead to their accurate reconstruction if one has access to many data points on both sides of the observation that one seeks. It is thus impossible to obtain good estimates at the edge of the sample.[12]

I also consider other, more specific, measures of accuracy. In particular, Table 2 gives the difference between the standard deviation of the estimated cycle and the standard deviation of the cycles that is embedded in the relevant artificial data series. Table 3 gives the difference between the first order serial correlation of the estimated cycle and the serial correlation of the true cycle. Lastly, Table 4 gives the coefficient β_v for $v = 10$. This gives an idea of the extent to which the estimated trend and cycle are related. This is of particular interest here because the artificial series contain uncorrelated trends and cycles.

One immediate conclusion from Tables 1-4 is that, as suggested by Baxter and King (1999), the results from using the HP filter and those from using the approximate band-pass filter they propose for keeping all frequencies whose period is longer than 32 quarters are very similar. What is somewhat more surprising, is that the band-pass filter proposed by Christiano and Fitzgerald for keeping the same frequencies leads to uniformly less accurate estimates of the true trend than those obtained by the HP method. This can be seen in Table 1 which shows that the MSE for the Christiano-Fitzgerald method is higher for every series. While the estimation of smooth trends is not the direct aim of Christiano and Fitzgerald (1999), it is nonetheless interesting that a method which is ideally suited for extracting low frequency variations under some circumstances does a systematically poorer job than the HP filter with my artificial data.[13] The one dimension in which of the Christiano-Fitzgerald "trend" is more attractive is that, as shown in table 4, expected movements in this trend are less associated with the cycle than in the HP case. On the other hand, Tables 2 and 3 show that the standard deviation and first order serial correlations of the Christiano-Fitzgerald "cycles" are smaller than those of HP filtered series. This is true even though the latter are smaller than the true standard deviations and first order serial correlations.

[12] As Baxter and King (1999) stress, the HP trend is subject to similar difficulties at the edge of the sample.

[13] This may be related to the fact that their method is optimized for situations where shocks that have long run effects also have large short run effects. By contrast, there are only trivial short run fluctuations associated with the long run movements in my artificial trends.

From now on, I thus focus attention on the polynomial trends, on the HP trend, and on the method that I propose. The linear detrending method proves to be adequate only when the true trend is linear. Otherwise, it has much higher MSE's than the trends that include quadratic and cubic powers of time. These latter trends have low MSE's when the trends are given by D^1, D^2, D^3 and D^6 with the quadratic trend being superior in this respect when the trends are given by D^2 and D^6 and the cubic trend being better when the trend is given by D^3. This latter effects is probably not surprising given how well a cubic trend captures the changes in D^3 itself. For these nearly-polynomial trends, the differences in MSE with my proposed method are quite slight. Indeed, when the true trend is linear, my method sometimes has a lower MSE than that from fitting a polynomial trend. This results, presumably, from sampling variation in the estimated coefficients of the high powers of time.

Interestingly, when the true trends are given by either D^1, D^2, D^3 or D^6 my proposed method tends to yield cycles whose standard deviation and serial correlation is slightly more accurate than those computed from polynomial trends. The differences do not appear large, however. More noteworthy differences between my proposed method and polynomial trends arise in the case of the more variable trends D^4, D^5 and D^7. In these cases, my proposed trend is more accurate than polynomial trends both in that its MSE is lower and in that it comes closer to reproducing the basic second moments of the cycle. This is true even though, as I mentioned earlier, a cubic trend can reproduce the path of D^4 with a great deal of accuracy.

I now turn to the comparison of my method with the HP filter. For both methods, the MSE depends on the nature of the cyclical series. As one moves from C^1 to C^2 and C^3, the MSE of the two trends rises. This may be because these cycles are increasingly variable, the variances of C^1, C^2 and C^3 equal 2.15, 7.62 and 9.96 respectively. This increased variability may well make it more difficult to obtain accurate trends. In the case of the HP filter, this problem may be compounded by the fact that, as shown in Table 4, its trend and cycle move together.

The MSE of the method I propose also depends on the nature of the true trend. It is higher for D^2 than for D^3 and, in turn, higher for D^3 than for D^1. By contrast, the accuracy of the HP trend is essentially unaffected by which of these three trends is employed.

As Tables 2 and 3 show, the HP trend leads to cycles that are both less variable and less serially correlated than the true ones in all the examples I consider.[14] Moreover, this reduction in the standard deviation and serial correlation of the cycle is always larger than that which results from the method I propose. On the other hand, this does not mean that the HP trend is always inferior. Indeed, there are two cases when my method leads to higher MSE's, namely when the trend equals D^5 so that it exhibits growth cycles and when the trend is given by D^7 while the cycle is an i.i.d. random variable. In these cases, my method produces cycles whose standard deviation and serial correlation are both excessively large. Moreover, the difference between the values of these statistics and the true values are larger in absolute value when using my proposed method than when using the HP filter.

For all other artificial series, my proposed method has a lower MSE and leads to more accurate estimates of both the standard deviation of the cycle and of its serial correlation. Broadly speaking, this suggests that the proposed method is more accurate than HP when the cycles are serially correlated and the true trends are not only smooth but also fail to contain cycles of high periodicity. The rather good performance of my proposed method in these cases - as well as their superiority from polynomial trends as the true trends deviate somewhat from being polynomial - suggests that one can indeed improve on existing methods. This is the key conclusion I draw from the analysis.

I now discuss some possible explanations for the relatively poor performance of my proposed method for y^{53} and y^{74}. I do this in part to suggest shortcomings that might be circumvented by future detrending methods. When the trend exhibits "growth cycles" my method finds it difficult to disentangle these growth cycles form the stationary component of

[14] The reason is, presumably, that the trend tracks short run movements somewhat. This may also explain why, in her simulated examples, Kozicki's (1999) multivariate detrending often leads to more accurate estimates of the serial correlation of the cycle. When using her multivariate HP procedure on series whose cycles differ by construction, the objective function has less to gain by letting the common trend track any one cycle.

the series. There is a sense, however, in which separating trends from cycles makes relatively less sense from the point of view of data description when the trend itself is subject to large short term fluctuations. The potential difficulties faced by my method in this case can be seen from a plot of the rates of change in the trend (which shows quite large changes). Also, β_v is more sensitive to v for this series. While Table 4 shows that β_{10} is still relatively small, β_{25} is equal to about .85. Perhaps one simply cannot expect to construct meaningfully independent trend and cycle series when the two have such large swings over relatively similar horizons. The result of this may be that the accuracy of the decomposition is not maximized by pursuing independence in the way I have been doing.[15]

The weakness of my method when the cycle is i.i.d. may well stem from my choice of $k = 16$. When the cycle is i.i.d., one can make the sharper prediction that cyclical observations that are much closer to one another should be orthogonal as well. Thus, the choice of $k = 16$ appears to makes sense only when one has *a priori* knowledge that cycles are relatively persistent.

It is worth noting that the fact that the HP filter out-performs my proposed method for the variable y^{74} is not surprising since this variable is constructed in such a way that the HP filter solves the optimal signal extraction problem for the trend. It may thus be more remarkable that when the trend remains equal to D^7 while the cycle is more persistent, the HP method does more poorly than the one I propose. This is true even though the relative variance of cycle and changes in trend growth is kept roughly the same. My proposed method

[15] If one reduces k to 12, β_5 is zero with λ equal to 4750 but this leads to an even higher value of β_{10}, namely .037. Thus, independence is reduced further. On the other hand, the mean square error neglecting the first and last 16 observations falls to 1.1. My conjecture for why accuracy can increase even though independence falls is the following. Suppose one conceives of the trend as a linear function of the observations, as in (5). If, in constructing the current observation of d, the weight of the current observation of y is significantly higher than the weight of any other observation, cyclical increases in y will be reflected in increases in d as well. Declines in the cycle following a boom will be correlated with declines in the trend. To make trend and cycle independent, the weight of the current y must not be too large relative to the weight of nearby y's. However, if the trend varies relatively fast, one can only get an accurate estimate of the current value of d by giving a great deal of weight to the current value of y. Thus, fast movements in the trend lead independence and accuracy to be conflicting objectives. By contrast, if the trend varies very slowly, accuracy is actually lost by giving a great deal of weight to current y because that induces too much similarity between the trend and the cycle.

does better, in part, because it leads to a larger values of the smoothing parameter λ as the cycle becomes more persistent. In particular, the λ for y^{74} is set at 15637, while those for y^{75} and y^{76} are set at 31819 and 34139 respectively. This makes intuitive sense. As the cycle has more power at lower frequencies, it makes sense to keep fewer low frequencies in the trend.

Before ending the comparison of my method with the HP trend, it makes sense to discuss the effect of these methods on the kinds of statistics for which HP filtering is often employed. As in King and Rebelo (1998), for example, the two types of statistics that are most commonly computed with HP filtered data are ratios of standard deviations and correlations. The question is then whether, or when, these statistics are sensitive to the use of the HP filter. Some of the relevant information is in Table 2. It shows that the distortion of the standard deviation of the cycle is much larger for series involving C^2 and C^3 than for series involving C^1. At the same time, Table 4 shows that the trend and the cycle are much more correlated in the case of C^2 and C^3. Because these two series decay more slowly, more of their power is at lower frequencies and, for this reason, more of their variance is incorporated in the HP trend.[16]

This can affect both ratios of standard deviations and correlations. In particular, the ratio of the standard deviations of the HP cycle of a series whose true cycle is C^1 to one whose true cycle is C^2 is larger than the ratio of the standard deviation of C^1 to the standard deviation of C^2. By contrast, the standard deviation of the series detrended with my method is essentially the same as the standard deviation of the cyclical components. This can be stated somewhat differently as saying that one would expect my method to produce a higher ratio of cyclical standard deviations than the HP method if the true cycle of the numerator has more power at low frequencies than that of the denominator.

Similarly, the correlation of detrended series containing C^3 and series containing C^1 is higher using the HP method than the corresponding true correlation between C^3 and C^1. The former is .61 while the latter is .52. The reason is that C^3 contains both C^1 and C^2,

[16] My method, instead, adjusts the parameter λ which determines the range of frequencies that enter into the trend. In particular, λ rises from 153,000 to 576,000 as one moves from detrending y^{21} to detrending y^{22}. By having more weight on a smooth trend, the method allows more low frequencies.

but the HP filter puts more of the fluctuations of C^2 in the trend. Thus, HP filtered series containing C^3 are relatively similar to HP filtered series containing C^1. For the same reason, the correlation between detrended series containing C^3 and series containing C^2 is lower using the HP method than the true correlation of these cyclical series. The former is .88 while the latter is .9. My proposed method, on the other hand recovers the correlation of the underlying cycle processes, whether the trend is D^1, D^2, D^3 or D^4.

What appears to matter here is the spectrum of the common series which induces correlation between the two variables. If this common series has relatively little power at low frequencies, the HP filtered series tend to be more strongly correlated than those detrended by the method I propose. If the common component has more power at low frequencies than the components which are not common, the HP filtered series tend to be more weakly correlated than those detrended by my method.

5 Differencing the Artificial Data

The analysis so far has focused on comparing the method of detrending I propose to other methods which are geared towards obtaining both relatively smooth trends and stationary cycles. An alternative approach for obtaining stationary time series from observations on the economy is to follow the approach advocated by Box and Jenkins (1970). This involves differencing time series until they appear stationary.

One advantage of this approach is that statistical analyses of differenced series also recover the effect of shocks on the long run level of a series. Estimation exercises thus yield answers to questions concerning both temporary and long term fluctuations. This is attractive because, as argued by Diebold (1998) it is often desirable to evaluate models by simultaneously studying their implications for long and short term movements. This is particularly true when, as in Rotemberg and Woodford (1996), one is evaluating models in which technological change has important cyclical consequences. Since technological change also has effects in the long run, studying whether the model can account for both long run and short run changes simultaneously seems more proper than looking only at its ability to account for cyclical changes.

A second purported advantage, and this is stressed by Box and Jenkins (1970), is that differencing allows for a variety of trends, including some very smooth ones. For example, the series containing the quadratic trend D^2 as well as D^7 can be represented as ARIMA processes that are I(2) while the series containing the trend D^6 follows an ARIMA process that is I(3). Even the series with a linear trend can be represented by an ARIMA process which is I(1).

What is unusual about these ARIMA processes, however, is that the MA polynomials of the resulting processes are noninvertible if the level of the series also contains a stationary component. In particular, the MA polynomials in these processes contain as many unit roots as the number of times one must difference the series to make it stationary. It has long been recognized that even a single unit root in the MA polynomial (which emerges when the series has a linear trend and a stationary disturbance) creates estimation problems. For this case, Plosser and Schwert (1977) showed by simulation that linear detrending leads to more accurate estimates than those one obtains by first differencing the series. In a more recent study that builds on a great deal of intervening literature, Davis and Dunsmuir (1996) compute the distribution of the maximum likelihood estimate of the MA coefficient that results from differencing a process with a linear time trend. They show that this estimate has greater variance around the true value of 1 than would be predicted on the basis of the standard asymptotic normal distribution. This is a more formal demonstration that differencing tends to yield inaccurate estimates.

These problems pale by comparison with those I encountered when I used my commercially available software (which does not use maximum likelihood) to fit series with quadratic trends by differencing the series twice.[17] In no case did this software detect that there were two unit roots in the MA polynomial. Indeed, the estimated process was in each case quite different from the actual one even though I was fitting the theoretically correct process.[18]

[17] I was unable to find a formal analysis of the case where the MA polynomial has more than one unit root.
[18] In the case of y^{21}, for example, this is an ARIMA(1,2,2).

Similarly, I did not estimate three unit roots in the MA polynomial when I differenced three times the series whose trend is given by D^6.

These issues are somewhat removed from actual practice since applied macroeconomists rarely fit MA components. This raises the question of what researchers who favor the differencing of nonstationary series would do if they wanted to analyze my artificial data. Perhaps a survey is necessary to answer this question adequately. Here, instead, I study these series in what I construe to be standard practice following Dickey and Fuller (1979) and Box and Jenkins (1970).

For each artificial series based on D^1, D^2, D^3 and D^6, I first carry out augmented Dickey-Fuller tests of the null hypothesis that there is a unit root against an estimated alternative in which the series is stationary around a linear trend.[19] When the trend is given by D^1, the augmented Dickey-Fuller tests correctly reject the null hypothesis. For the series whose trends are given by D^2, D^3 and D^6, the top of Table 4 displays the augmented Dickey-Fuller regressions. In each case, one cannot reject the presence of unit root.[20]

The bottom of Table 5 gives my estimates for stochastic processes followed by the first difference of these series. The approach to obtaining these is based on Box-Jenkins (1970). In particular, the portmanteau Q statistic for misspecification does not detect any abnormal autocorrelations. I also differenced the series a second time and carried out augmented Dickey-Fuller tests for the presence of a second unit root. These rejected this hypothesis with a high degree of confidence (even when an AR(x,2,2) with x equal to one or two would correctly describe the actual process). My conclusion is that the widespread tendency to accept the existence of a unit root when the hypothesis cannot be rejected, as well as the parsimony of the resulting representations would lead researchers to conclude that these series have a unit root.

[19] I obtained similar results assuming that the alternative does not include a linear trend.

[20] The t-statistic against the hypothesis that the lagged level variable is insignificantly different from zero is substantially larger when the square of the trend is included in the regression and the true trend is D^2. Presumably, this increase in the t-statistic is enough to reject the null hypothesis of a unit root in this case. However, special tabulations are probably necessary to make sure of this since tests of this null hypothesis are sensitive to the inclusion of trending variables. It is also worth mentioning that this t-statistic also rises substantially when a square trend is included in regressions explaining the log of U.S. GDP.

Shocks to the series would then be deemed to have permanent effects even though, as we saw, the effect of shocks to C^1 essentially disappear after 4 years. The hypothesis that these series have a unit root would also tend to lead one to study bivariate relationships by running a VAR where the variables enter in first differences. This is particularly true if one is building a statistical model of y^{21} and y^{33} since the difference between these variables changes over time so that they are surely not cointegrated. I have thus studied such a VAR, together with the corresponding VARs for the series detrended by setting k equal to 16 and the VAR for the true cyclical series C^1 and C^3.[21]

Since C^3 consists of C^1 and another random variable, it is appealing to run the VAR with the true cyclical series by letting C^3 be affected contemporaneously by C^1 while letting C^1 be affected by C^3 only with a lag. Including two lags of each cyclical series then leads to a simple VAR which recovers the true processes for the underlying series. The shock to C_t^3 which is orthogonal to the shock to C_t^1 is simply the shock to C_t^2, ϵ_t^2. This shock thus leads to a hump shaped response of C^3 and no response of C^1. By contrast, the shock to C_t^1, ϵ_t^1 leads C^1 and C^3 to respond in the same way. To conserve on space, I thus report only the response of C^3 to the two shocks in Figure 8.

I fit the same VAR to the series c_t^1 and c_t^3 that are obtained from y_t^{21} and y_t^{33} respectively by detrending them with the method I propose. The figure also shows the responses of c_t^3 to one-standard-error shocks of the two series. These are essentially identical to the responses of the true underlying cyclical series. This, too, is not surprising given the ability of my method to recover the true cyclical series in these cases.

Finally, I fit a VAR to the first differences of y_t^{21} and y_t^{33}. It is not completely clear how many lags ought to be included in the VAR but, because the third lag is statistically insignificant, I also include two lags in this VAR. By putting y_t^{21} first, I ensure once again that there is no response of y_t^{21} to the shocks of y_t^{33}. Figure 8 then gives the resulting responses of y_t^{33} to a one standard error response in the two series that I include in the VAR. The

[21] To avoid the effect of the observations near the boundary of the sample and to preserve comparability, all these VAR's are run dropping the first and last 16 observations.

response of the level of this series is obtained by integrating forward the estimated responses of the first differences of y^{33}.

The initial estimated response to the two one-standard-error shocks is essentially the same as the actual response. The subsequent responses are widely off the mark, however. First of all, both shocks are now estimated to have permanent effects on y^{33}. More surprisingly, the pattern of the response of y^{33} to a shock to its own fist difference is wrong. The impulse response shows this series rising continuously after this shock, with fairly pronounced rises 3 and 4 quarters after the shock. By contrast, the underlying series actually declines 3 and 4 quarters after such a shock takes place.

That short run responses are distorted by differencing may be surprising because the differencing filter is thought to emphasize high frequencies. One might thus expect it to lead to correct inferences about high frequencies. It is important to remember, however, that when a series is not stationary, the differencing filter can also accentuate some very slow-moving components. Suppose, for example, that the trend rate of growth of two time series are very different at the beginning and at the end of the sample. For purposes of illustration, suppose that these secular differences in growth rates are much larger than the typical changes in growth rates from one period to the next. The regression coefficient of one differenced series on another will then be influenced mainly by the connection between the secular growth rates. This problem may be of some practical importance, as the average rate of growth of GDP is lower after 1973 than before.[22]

An argument that is often made against detrending (see Harvey (1979) for instance) is that this leads to a distorted picture of the behavior of series that truly have unit roots. This is obviously correct and, indeed, detrending random walks using either my method or the HP filter leads to stationary series whose behavior is quite differently from that of the true underlying process. My point in this section is that quite poor inferences can also be made if a series whose trend is smooth is differenced before being subject to statistical analysis.

[22]In this regard, it is interesting and somewhat disturbing that, in spite of these changes in secular growth rates, one hardly ever finds that macroeconomic time series have two unit roots (i.e. a unit root in their growth rates).

Moreover, these poor inferences seem to come about precisely because of what I first termed an advantage of the approach that proceeds by differencing time series. With this approach, it appears to be possible to answer questions concerning both the short and the long run effect of shocks. However, if the specification that one ends up with is incorrect (perhaps because of the difficulties caused by noninvertible MA polynomials) the short run and long run properties of time series end up commingled in estimation in such a way that misleading conclusions are reached about both.

It thus seems worthwhile to develop alternative estimation strategies for the case of smooth trends, and this paper is a step in this direction. This is not the end of the story, however. One would then wish to test the hypothesis that the trend is indeed smooth against the specifications that are generated by the approach based on Box and Jenkins (1970). This testing task still lies ahead.

6 Application to Aggregate Data

The previous sections have motivated the use of my detrending method with $k = 16$ and $v = 5$ by arguing that, in certain circumstances, it uncovers the properties of the underlying data generating process when other approaches do not. This provides one motivation for detrending actual time series by this method. In addition, smooth trends are easy to describe because they involve only small changes in their rate of growth and, often, these changes have the same direction for long periods. There is thus some hope that smooth trends which respond relatively flexibly to the actual evolution of time series contain useful information about the sources of secular changes.

In this section, I compare the properties of the cycles that result from applying the HP filter and my proposed method to some standard time series. I concentrate on the HP filter both because it is widely used and because the mean values for the squared second difference of the trends constructed using the Baxter-King and Christiano-Fitzgerald methods are larger. Thus, if ones aim is to construct smooth trends, the HP method is preferable. One of the conclusions of this section, however, is that one obtains considerably smoother trends by applying the method proposed in this paper. At the same time, the cycles are short lived in either case.

As I showed in Section 4, the HP filter and my proposed method handle series with relatively persistent temporary components quite differently. This makes it of interest to study military purchases. U.S. government purchases for national defense are depicted in Figure 9. This Figure shows rather clearly that these increased in the Korean War, the Vietnam War and during the "Reagan buildup". Except for the Korean War, these increases were fairly protracted. The HP filter may thus include a high fraction of these changes in the trend so that the filtered data could mask the "short run" association between these defense buildups and other variables. If my proposed method includes more of these short run movements in its cycle, it may thus render visible some effects of military buildups. It is thus particularly interesting to study how other variables are related to military purchases using the two approaches.

The quarterly post-War series that I analyze are thus the log of U.S. GDP, consumer expenditure on nondurables and services, investment, purchases for national defense, total hours in the nonfarm sector, the ratio of GDP to the hours in the nonfarm sector (as a measure of productivity), the ratio of labor compensation deflated by the GDP price deflator divided by this measure of hours (as a measure of the real product wage), the labor share constructed by dividing nominal labor compensation by nominal GDP, industrial production, the unemployment rate and the log of Japanese GDP. For most series I use data from 47:1 to 98:1, though unemployment starts in 1948 and Japanese GDP starts in 1957.

My reason for studying the unemployment rate is that it is often treated as *the* cyclical variable in empirical studies in spite of the appearance of at least some local trends. While Blanchard and Quah (1989) deal with this apparent nonstationarity by linearly detrending this variable in their study, Shimer (1998) shows that the medium term changes in unemployment caused by demographic changes are substantially more complicated. It thus seems worthwhile to see if detrending leads to a tighter association of unemployment with other measures of the business cycle.

I include Japanese GDP, from the IMF International Financial Statistics, because some of its properties seem to hinge on the detrending method that is employed. While Taylor (1989) has argued that Japanese GDP is less cyclically variable than U.S. GDP (and has provided a model to explain this), the standard deviation of Japanese GDP growth exceeds the standard deviation of U.S. GDP growth over the same period. It thus seems worthwhile to see what the HP filter and my method say about their relative variabilities (and their correlation).

Table 6 displays the values of the smoothing parameter λ which ensures that (3) holds. The highest of these estimates is less than 4 times higher than the smallest one. More importantly, the λ's for hours, GDP and productivity are even more tightly clustered (near 70,000). This means that the difference between detrended log output and detrended log hours is essentially identical to detrended log productivity.

Because of my desire to keep the observations around the Korean War in the sample I analyze, the statistics in this Table are constructed with the full sample of observations available for each series. Relative to dropping the first and the last 16 observations, this leads to a stronger association of military purchases with other variables. Otherwise, the use of the full sample does not have really substantial effects on the statistics reported in the Table. One clear difference between the HP method and the one I propose is that the latter are smoother. The mean value of the first difference in trend growth is always larger for the HP trend, sometimes considerably so. The ratio of these mean values goes from 4.9, in the case of Japanese GDP to 132, in the case of U.S. investment. For U.S. GDP, the ratio is 51, which is close to the average of the ratios for the variables in the table.

Consistent with the notion that the HP method tends to make cycles return relatively quickly towards the mean, the table shows that the first order serial correlation of the cycle is always higher when $k = 16$. On the other hand, both methods lead to cycles that are effectively over after four years as can be seen from the correlations of the cycles at t with the cycles at $t - 16$ displayed in the Table. These correlations are fairly small in absolute value and mostly negative, particularly for my proposed method. In the case of my proposed method, these statistics are somewhat affected by the inclusion of observations near the boundaries of the sample. When ignoring the first and last 16 observations, the average of these serial correlation statistics equals -.1 for both methods. When using the full sample, by contrast, the average of these statistics is -.21 for my proposed method while it equals -.09 for the HP method.

The Table also shows that many of the business cycle facts emphasized in the real business cycle literature do not hinge on whether the HP filter or my proposed method are used. In particular, both the relative standard deviation and the correlation with GDP are essentially the same for consumption, investment, hours, productivity, unemployment and industrial production. While the difference is small, the correlation of cyclical output and cyclical unemployment is higher with the proposed method. Such a high correlation seems desirable if one wishes to interpret cyclical unemployment as a good indicator of the business cycle in general.[23]

The biggest percentile difference in the relative standard deviations of the cycle produced by the two methods concerns Japanese GDP. The contrast that I find may be due in part to the fact that the HP trend is more strongly associated with the HP cycle in Japan than in the U.S. In particular, my measure of association β_5 equals .27 in Japan and only .18 in the U.S. This suggests that the HP trend in Japan captures some of the fluctuations that I treat as cyclical.

In absolute value terms, the two variables whose cyclical correlation with GDP are most affected by the detrending method are military purchases and wages. The difference in the correlation of the labor share with GDP is presumably explained by the difference in the real wage correlation since both real wages and the labor share are more procyclical with my proposed method. These differences suggest that the components of real wages and output that revert to their smooth trends relatively slowly are more strongly correlated with one another than the movements that are extremely transitory. My proposed method also

[23]It is interesting in this regard that the correlation of cyclical unemployment with cyclical output I compute is higher than the correlation of the change in quarterly GDP with the change in unemployment. This "Okun's law" correlation equals .71.

indicates a stronger positive association of military purchases and output.[24] Given that changes in military purchases are relatively long-lived, the fact that this is larger with my method suggests that the effects on output of changes in defense purchases are relatively long-lived as well.

I now consider the connection between cyclical movements in the real product wage and in purchases for national defense. Using the HP filter, the correlation between these variables is .28 while, using my method, it is a remarkable .7.[25] Figure 10 shows both detrended real wages and detrended military purchases. Detrended military purchases are very similar to actual purchases, in that they have three big bulges. What is remarkable about the Figure is that real product wages also exhibit just three extended periods where they are above trend and these are roughly coincident in time with the bulges in military purchases. The main exception is the bulge of wages which peaks in 1974. While this starts with the Vietnam buildup, wages continue rising even as defense purchases fall after 1969.

The connection between these variables has been the subject of some controversy. Rotemberg and Woodford (1992) showed that changes in military purchases were associated with increases in real product wages, which is consistent with the theory they present.[26] Using revised data, Ramey and Shapiro (1998) showed that real wages actually declined after the three dates (1950:3, 1965:1 and 1980:1) that mark the beginning of the hostilities that preceded these three military buildups. Their procedure involves using a piecewise linear trend whose slope changes discontinuously in 1973:2 and running a regression of wages on eight lags of a dummy that takes a value of one at the beginning of hostilities.[27]

Interestingly, wages detrended with my proposed procedure are not particularly high in the immediate aftermath of the dates that Ramey and Shapiro (1998) report for the

[24] If one drops the first and last 16 observations, the correlation of cyclical military purchases with cyclical output falls from .21 to .12 using the HP filter and from .49 to .39 using the method I propose.

[25] If, instead, the first and last 16 observations are dropped, this falls to .52. The correlation is also somewhat lower is one uses the ratio of hourly compensation in the business sector over the deflator for the business sector. It is then around .4.

[26] The theory has less clear predictions for the wage divided by prices that include rents on housing and the prices of imported raw materials. For this reason, I focus on real product wages.

[27] Similar results are reported by Edelberg, Eichenbaum and Fisher (1998) for VAR's in levels.

beginning of hostilities. The correlations between my measure of detrended wages and the current value as well as 5 lags of their dummy variable is slightly negative so that regressions of the sort they utilize do not reveal the strong association between military purchases and wages that is visible in Figure 10. By the same token, if one detrends wages using a piecewise linear trend with a discontinuity in trend growth in 1973:2, the level of detrended wages remains strongly positively associated with the level of military purchases. Thus, the sign of the correlation between wages and military purchases does not seem very sensitive to which of these detrending methods is used. Rather, it seems to depend crucially on whether one looks at wages right after hostilities abroad begin or whether one looks at them when national defense purchases are in fact high.

Before closing this section, I briefly take up the question of the stability of my proposed method. Unlike linear filters such as the HP filter, this procedure has the potential for being unstable with respect to changes in sample size, as the parameter λ need not stay constant when the sample is expanded and the procedure is applied anew. A more thorough study of the severity of this problem is left for further research. For many of the series considered in this section, however, the magnitude of the problem appears to be somewhat limited because the estimate of λ is not overly sensitive to the period of estimation. In the case of U.S. GDP, for example, the λ that satisfies (3) when one uses only the observations until 79:4 is 80682. This is only about 15% larger than the λ obtained in the full sample. The estimates of the trend obtained from these two values of λ are quite close to one another as well. The maximum difference between these trends, which occurs in the first observation of the sample, is .003. Ignoring the first and last 16 observations, the difference in trend values is always smaller than .001.

There are, in addition, four other variables for which the λ's estimated over these two subsamples are similar in that the larger is no greater than 1.5 times the smaller one. In other cases the λ's estimated over different samples are further apart. The labor share represents an extreme case; it is estimated to have an arbitrarily high λ (which I set at $8e^{11}$) until 1979:4 while the λ for the whole sample is only about 98,000. If one uses these two values of λ over the whole sample, the resulting trends differ by up to .03 (at the end of the sample). Even in the middle of the sample, differences just below .02 are not uncommon. This instability is troubling, though the relatively abrupt change in the trend growth of this series around the end of the 1970's causes similar difficulties for polynomial trends. If cubic trends are fitted over the two subsamples, the resulting trends are closer to one another in the middle of the sample but differ by .06 at the end of the sample.

7 Conclusions

In this paper I have provided a method for detrending time series by separating them into a smooth component and a stationary stochastic component. The method has some intuitive appeal and appears to do a good job of reconstructing such components when it is applied to some artificial time series which actually contain smooth trends. Still, I have not provided any general demonstration of its properties. All I have shown is that, there are plausibly important sets of circumstances in which it is possible to do considerably better than with existing methods. This suggests that further research on this topic is likely to be fruitful.

I hope to have also shown that the use of smooth trends provides some interesting insights into the behavior of macroeconomic time series. For one thing, it simplifies the description of the nonstationary trends by having these be subject to only small changes in their rate of growth. In practical applications, these changes in the rate of growth seem to have the same sign for long periods of time, and this may help facilitate discussion about relatively slow-moving forces that lead to secular changes. The second descriptive advantage of using a very smooth trend is that the series detrended by this method appear to capture much better the effects of relatively long-lived shocks such as changes in military purchases.

A final benefit of studying what happens when one allows for very smooth trends is that it should allow for a better understanding of the extent to which the empirical findings obtained using data in first differences are robust to reasonable alternatives. In the examples I present, the impulse responses from a VAR using differenced data are radically different from those of the true data generating process (which are uncovered by using the method I propose). This fits with the big difference in the impulse responses obtained by Blanchard and Perotti (1999) depending on whether they use quadratic trends or first differences. When such differences arise in practice, one needs good tests of the two models of trends against one another. Unfortunately, standard Dickey-Fuller tests seem unable to detect the falsehood of differenced stationary models in the simple examples I present. A final aim of this paper is thus to stimulate the search for better tests of this kind.

8 References

Baxter, Marianne and Robert G. King, "Measuring Business cycles: Approximate Band-Pass Filters for Economic Time Series," *International Economic Review*, 1999.

Blanchard, Olivier J. and Roberto Perotti, "An Empirical Investigation of the Dynamic Effects of Changes in Government Spending and Taxes on Output," mimeo, 1999.

Blanchard, Olivier J. and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4), September 1989, 655-73.

Bowley, Arthur L. *Elements of Statistics*, London: King and Sons, 1920.

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

Burns, Arthur F. and Wesley C. Mitchell, *Measuring Business Cycles*, New York: National Bureau of Economic Research, 1946.

Canova, Fabio, "Detrending and Business Cycle Facts," *Journal of Monetary Economics*, 41, 1998, 475-512.

Cogley, Timothy and James M. Nason, "Effects of Hodrick-Prescott Filter on Trend and Difference Stationary Time Series," *Journal of Economic Dynamics and Control*, 19, 1995, 253-78.

Davis, Richard A. and William T. M. Dunsmuir, "Maximum Likelihood Estimation for MA(1) Processes with a Root on or Near the Unit Circle," *Econometric Theory*, 12 (1), March 1996, 1-29.

Dickey, David A. and Wayne A. Fuller, "Distribution of the Estimators for Autoregressive

Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 1979, 427-31.

Diebold, Francis X., "The Past, Present and Future of Macroeconomic Forecasting," *Journal of Economic Perspectives*, 12(2), Spring 1998, 175-92.

Eichenbaum, Martin and Jonas D. M. Fisher, "How does an Increase in Government Purchases Affect the Economy," *Economic Perspectives*, 22(3), 1998, 29-43.

Ellison, Glenn and Drew Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy*, 101(1), August 1993, 612-43.

Hamilton, James D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," ; *Econometrica*, 57(2), March 1989, 357-84.

Hodrick, Robert J, and Edward Prescott, "Post-War Business Cycles: An Empirical Investigation,", *Journal of Money, Credit and Banking*, 29 (1), February 1997, 1-16.

King, Robert G. and Sergio T. Rebelo, "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control,* 17 (1), January 1993, 207-231.

— and —, "Resuscitating Real Business Cycles," forthcoming in John B. Taylor and Michael Woodford, eds. *Handbook of Macroeconomics*, Amsterdam: North Holland, 1998.

Kozicki, Sharon, "Multivariate Detrending Under Common Trend Restrictions: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control*, 23(7), June 1999, 997-1028.

McCallum, Bennett T. and Edward Nelson. An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis, Cambridge, MA: NBER Working Paper 5875, 1997.

Malinvaud, Edmond, *Statistical Methods of Econometrics*, Amsterdam: North Holland, 1980.

Mansfield, Edwin, *The Economics of Technological Change*, New York: Norton, 1968.

Nerlove, Marc, David M. Grether and Jose L. Carvalho, *Analysis of Economic Time Series: A Synthesis*, New York: Academic Press, 1979

Perron , Pierre, "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica*, 57(6), November 1989, 1361-403.

Plosser, Charles I. and G. William Schwert, "Estimation of a Non-Invertible Moving Average Process; The Case of Overdifferencing," *Journal of Econometrics*, 6(2), Sept. 1977, 1 99-224.

Ramey, Valerie and Matthew Shapiro, "Costly Capital Reallocation and the Effects of Government Spending," *Carnegie-Rochester Conference Series on Public Policy,*, 48, June 1998, 145-94.

Rotemberg, Julio J. and Michael Woodford, "Oligopolistic pricing and the Effects of Aggregate Demand on Economic Activity,' *Journal of Political Economy*, 110(6), 1992, 1153-1207.

— and — "Real Business Cycle Models and the Forecastable Movements in Output, Hours and Consumption," *American Economics Review*, 86, 1996, 71-89.

Shimer, Robert, "Why is the U.S. Unemployment Rate So Much Lower?," forthcoming in *NBER Macroeconomics Annual*, 1998.

Taylor, John B., "Differences in Economic Fluctuations in Japan and the United States: The Role of Nominal Rigidities, *Journal of the Japanese and International Economies*, 3(2), June 1989, 127-44.

Winston, Wayne *Operations Research: Applications and Algorithms, 3d edition*, Belmont, CA: Wadsworth, 1994.

Figure 1
U.S. GDP and its
Linear Trend

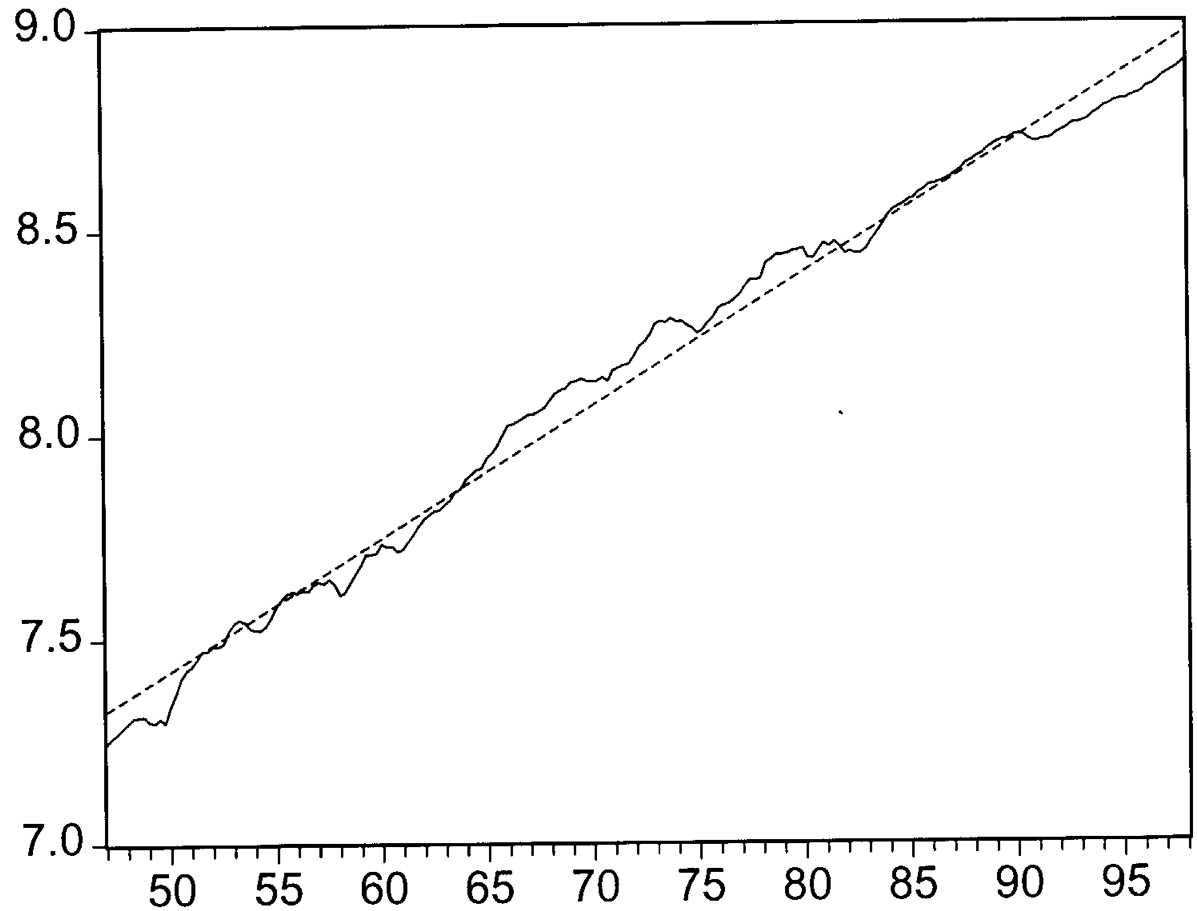

Figure 2
U.S. GDP and Two
Cubic Trends

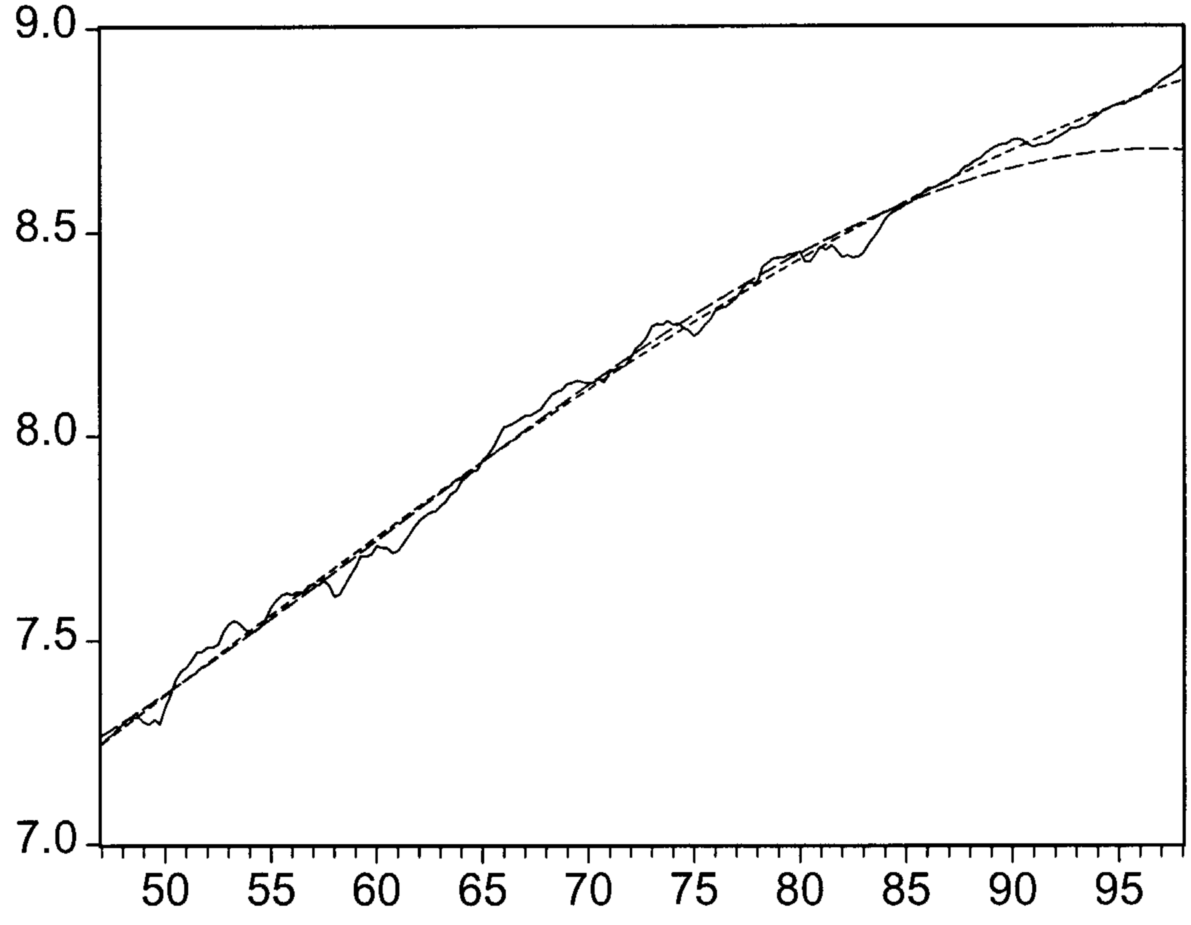

Figure 3
HP Trend Growth Rate
of U.S. GDP

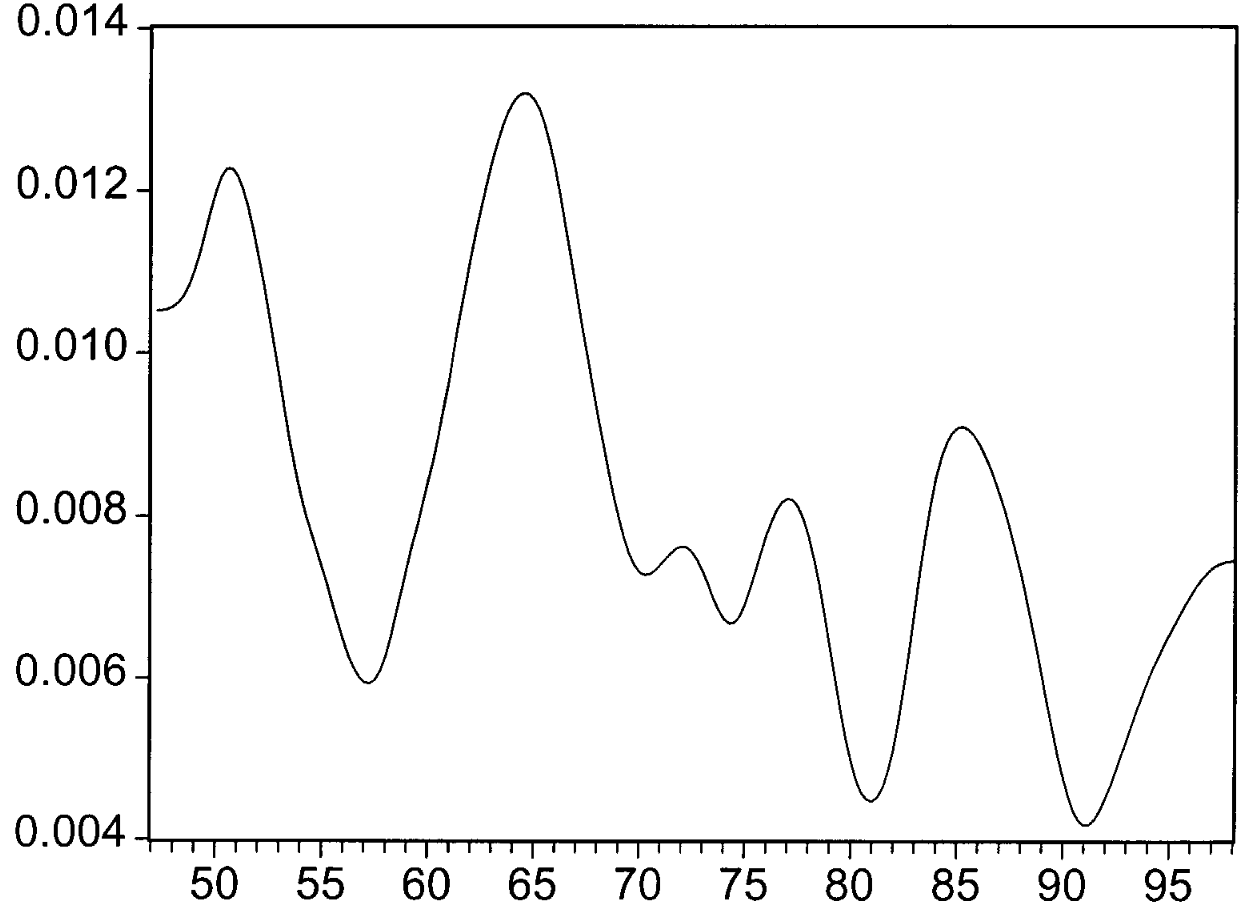

Figure 4
Trend Growth of U.S GDP
According to Proposed method

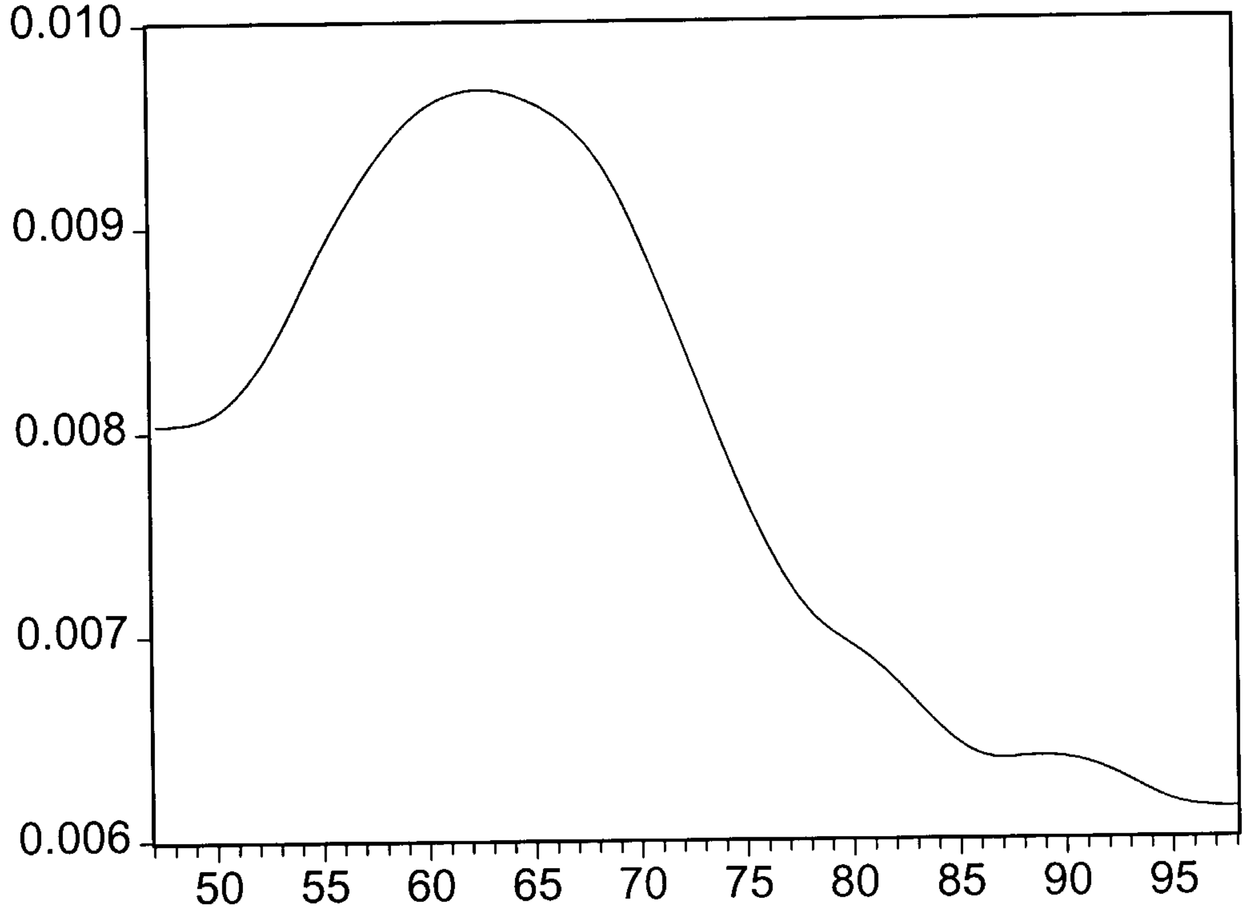

Figure 5: Gains for Hodrick-Prescott and k=16 with lambda=70000

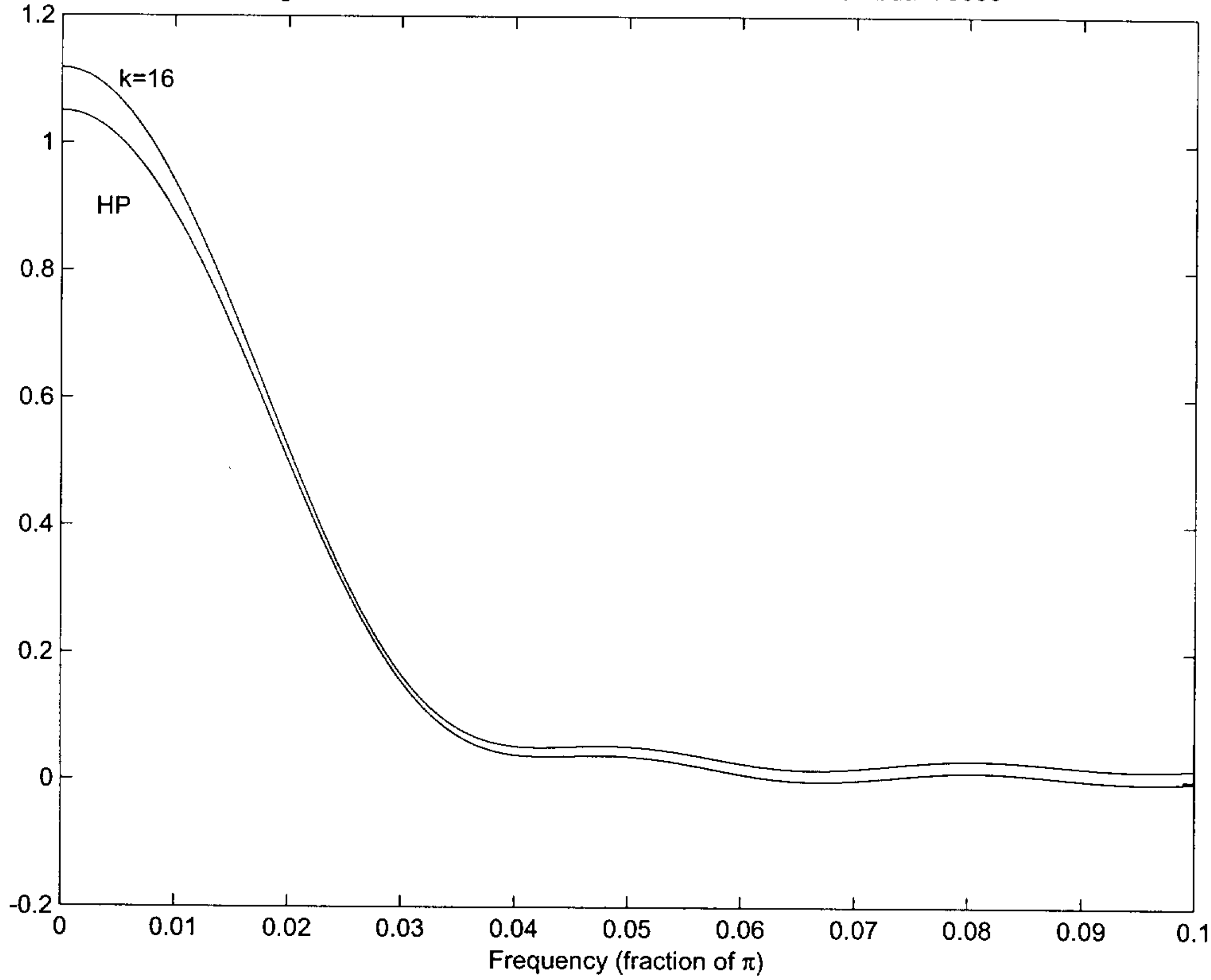

Figure 6
Artificial trend D^4

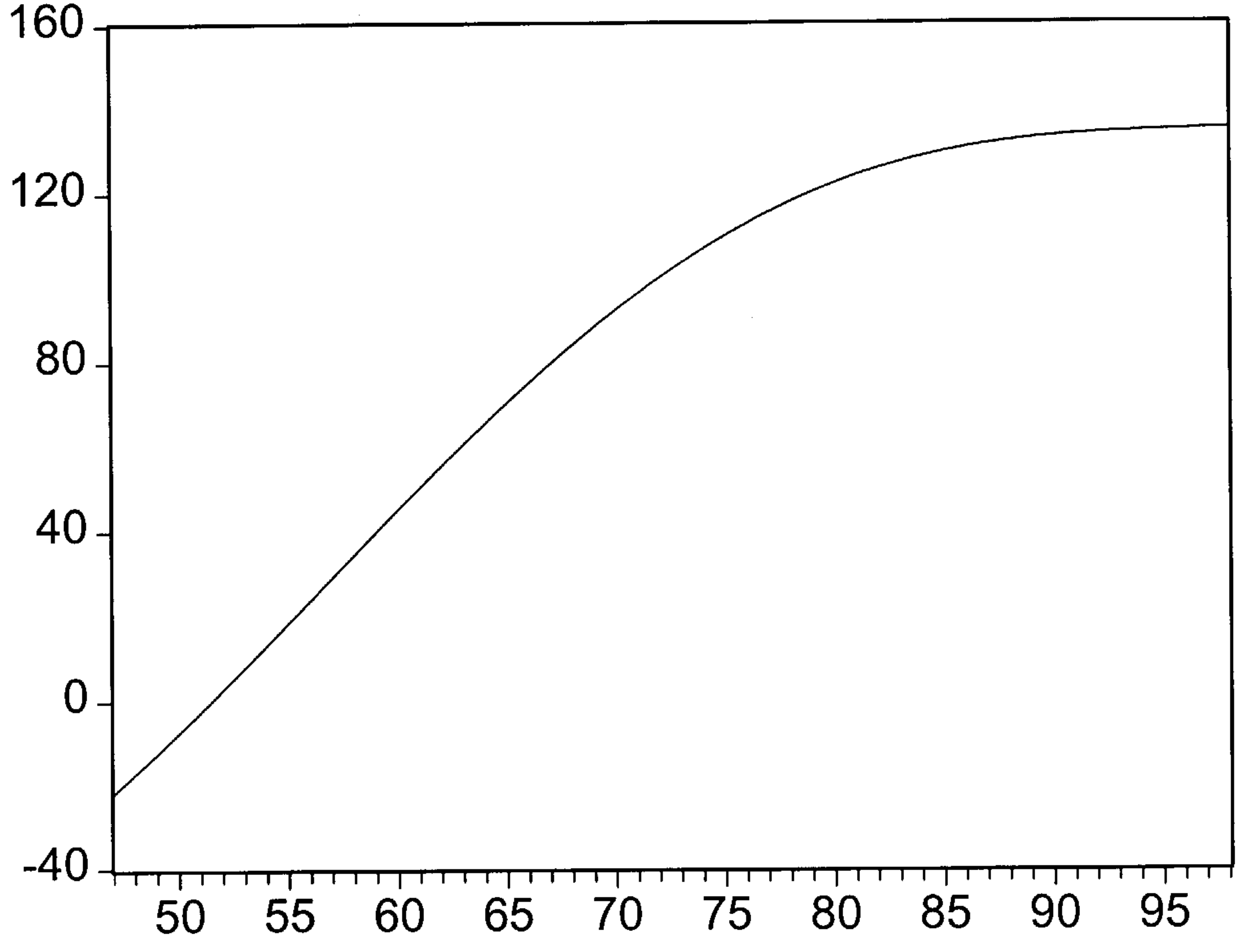

Figure 7
Artificial trend D^6

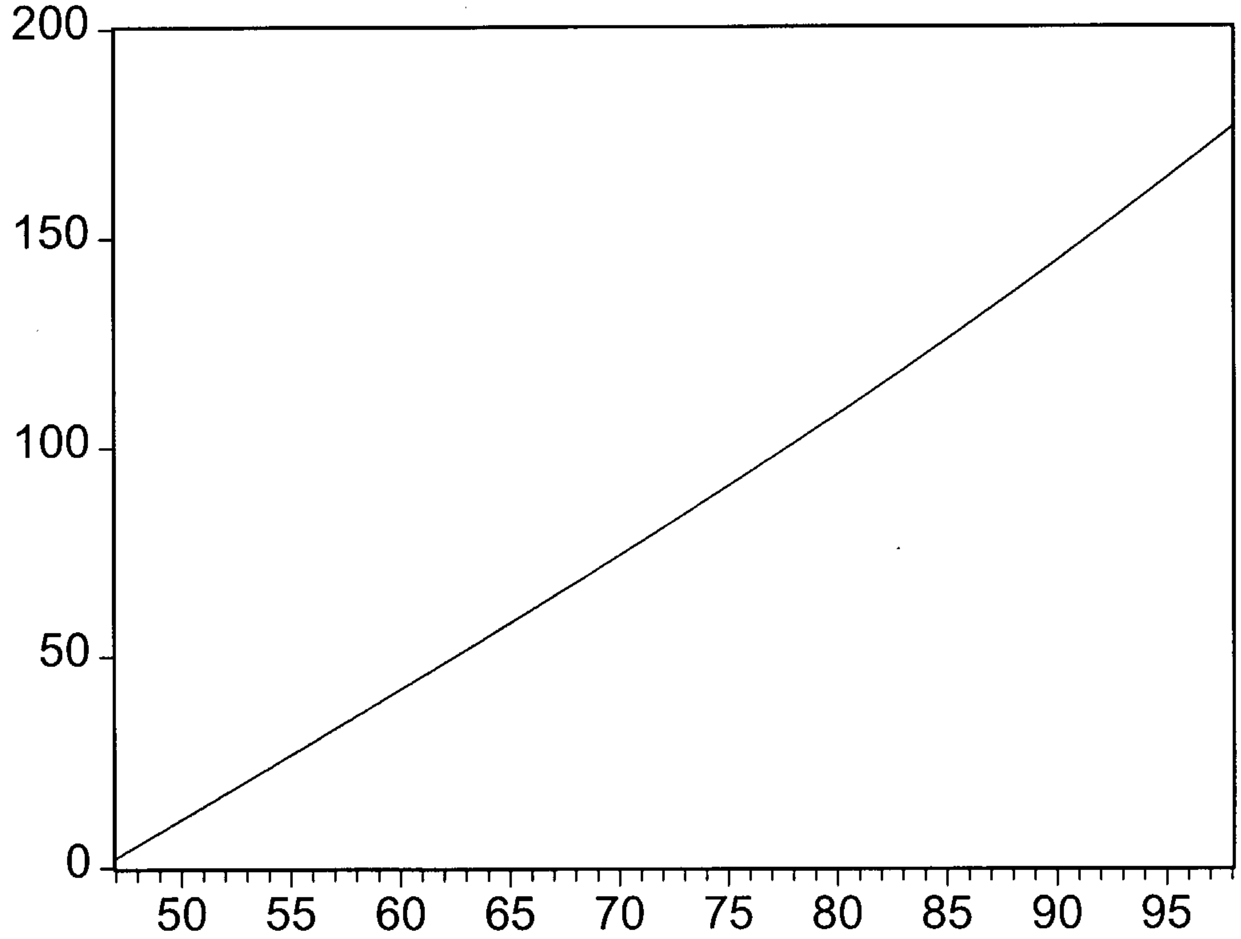

Figure 8: Response of cyclical y^{33} to shock of cyclical y^{21}

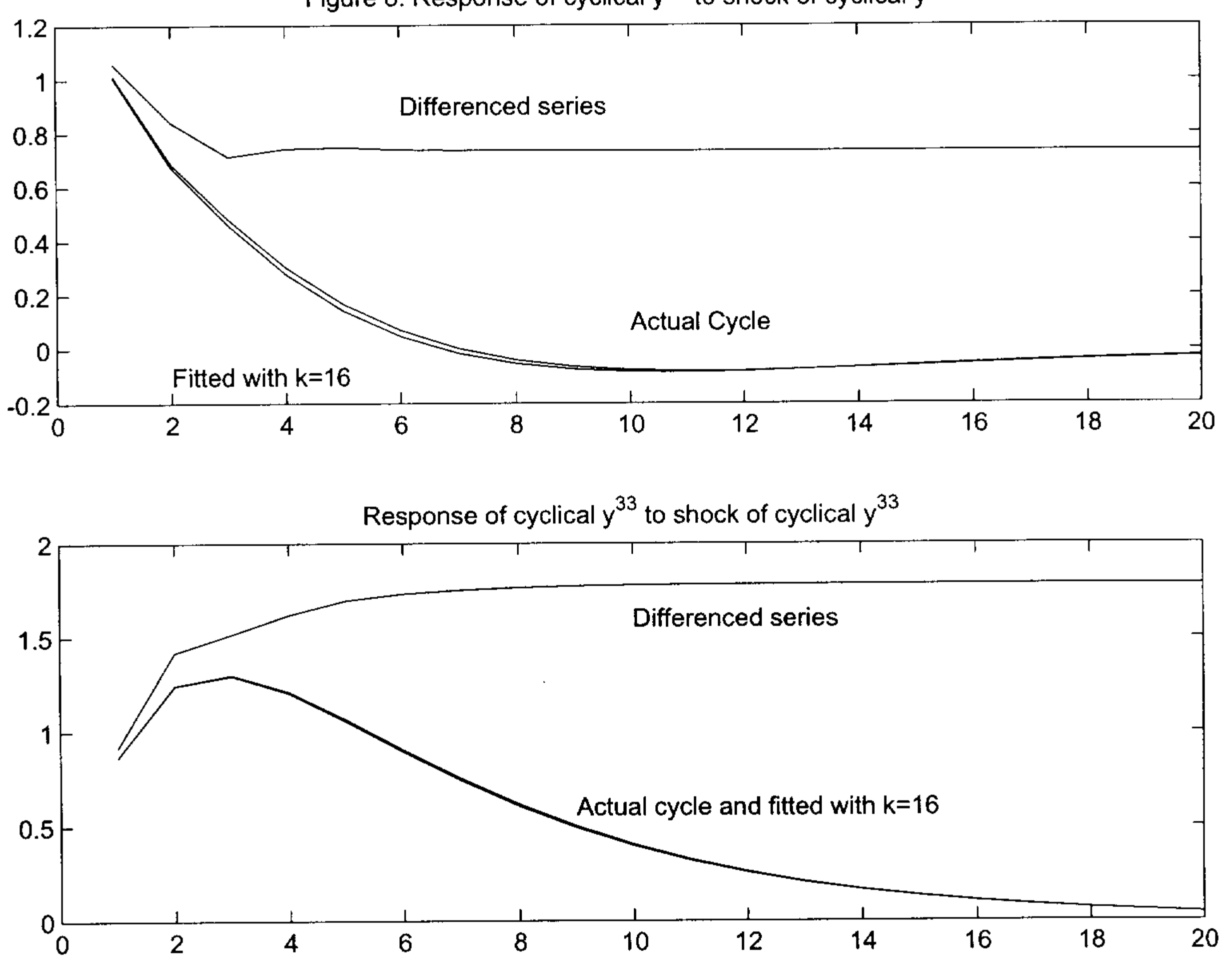

Figure 9
U.S. real purchases for
national defense

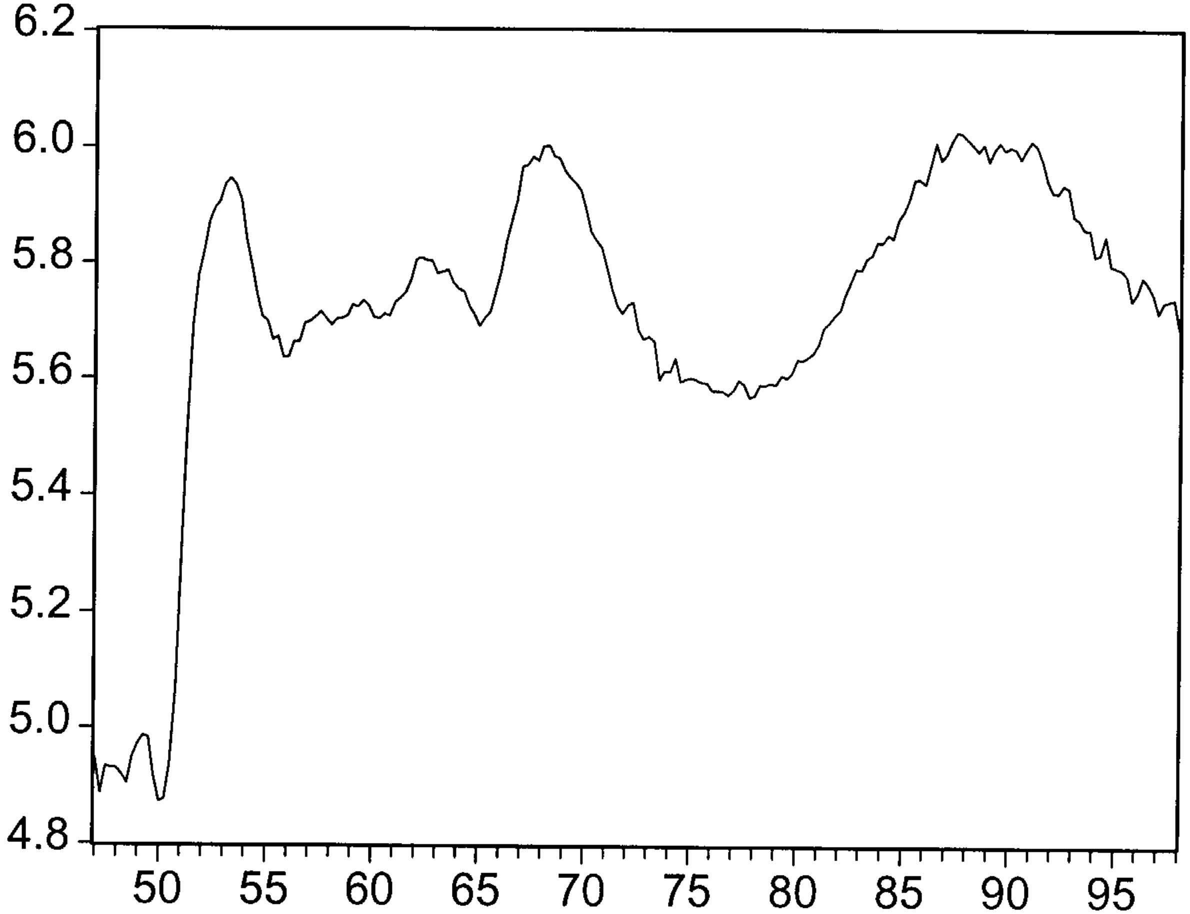

Figure 10
Detrended real wages
and national defense

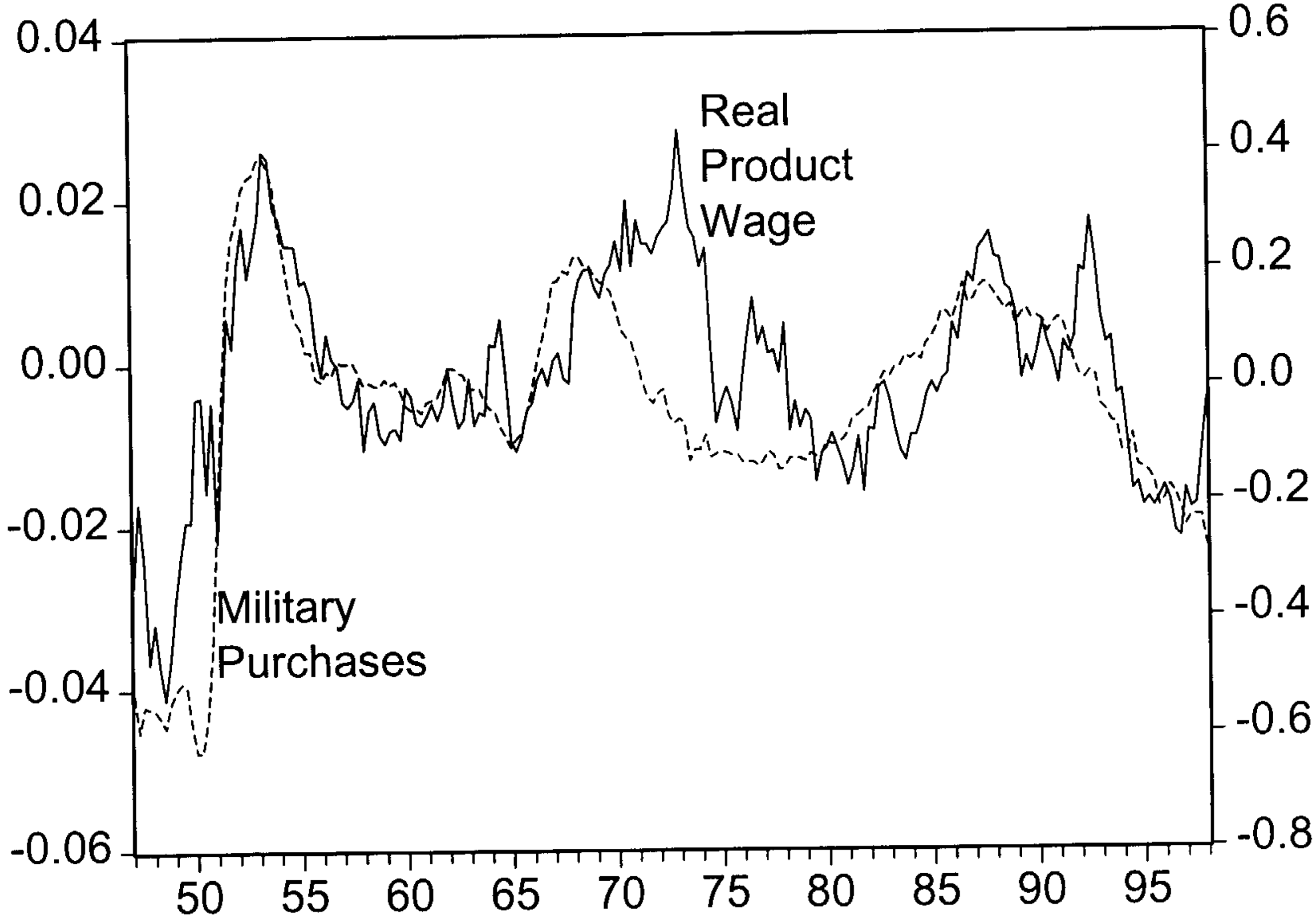

Table 1
Mean Square Errors

Var.	Trend/cycle		Detrending method						
			Linear	Quadratic	Cubic	Hodrick-Prescott	Baxter-King	Christiano-Fitzgerald	Proposed k=16
y^{11}	D^1	AR(1)	0.01	0.02	0.10	0.35	0.43	0.69	0.02
y^{12}	"	AR(2)	0.17	0.18	0.12	2.12	2.53	4.92	0.13
y^{13}	"	Sum	0.23	0.26	0.30	2.17	2.80	5.60	0.22
y^{21}	D^2	AR(1)	13.64	0.02	0.10	0.34	0.43	0.70	0.13
y^{22}	"	AR(2)	13.52	0.18	0.12	2.12	2.51	4.94	0.79
y^{23}	"	Sum	13.30	0.26	0.30	2.17	2.79	5.63	1.01
y^{31}	D^3	AR(1)	19.57	0.26	0.10	0.35	0.45	0.68	0.26
y^{32}	"	AR(2)	20.25	0.62	0.11	2.13	2.56	4.92	0.25
y^{33}	"	Sum	20.67	0.73	0.29	2.18	2.85	5.60	0.46
y^{41}	D^4	AR(1)	142.92	3.83	2.02	0.35	0.47	0.71	0.24
y^{42}	"	AR(2)	142.90	4.71	2.14	2.13	2.60	4.99	0.68
y^{43}	"	Sum	142.21	5.10	2.59	2.18	2.89	5.69	1.09
y^{53}	D^5	Sum	157.19	150.80	55.16	2.17	3.09	5.47	4.34
y^{63}	D^6	Sum	3.18	0.20	0.31	2.17	2.80	5.58	0.23
y^{74}	D^7	i.i.d.	45018.31	1747.83	1805.19	56.01	54.40	108.76	244.75
y^{75}	"	AR(1)	44751.51	1730.48	1872.23	387.71	374.45	602.11	292.90
y^{76}	"	AR(2)	44697.15	1719.02	1881.04	486.22	463.08	722.25	327.53

Table 2
Relative Standard Deviation of the Cycle

Var.	Trend/cycle		Detrending method						
			Linear	Quadratic	Cubic	Hodrick-Prescott	Baxter-King	Christiano-Fitzgerald	Proposed k=16
y^{11}	D^1	AR(1)	-0.003	-0.004	-0.033	-0.292	-0.325	-0.322	-0.020
y^{12}	"	AR(2)	-0.013	-0.029	-0.039	-0.899	-1.023	-1.129	-0.021
y^{13}	"	Sum	-0.029	-0.053	-0.054	-0.887	-1.039	-1.135	-0.029
y^{21}	D^2	AR(1)	2.188	-0.004	-0.033	-0.292	-0.325	-0.338	-0.032
y^{22}	"	AR(2)	1.312	-0.029	-0.039	-0.900	-1.023	-1.130	-0.002
y^{23}	"	Sum	1.115	-0.053	-0.054	-0.887	-1.039	-1.143	-0.022
y^{31}	D^3	AR(1)	2.829	-0.025	-0.036	-0.293	-0.324	-0.302	0.017
y^{32}	"	AR(2)	2.362	0.049	-0.037	-0.898	-1.022	-1.129	-0.027
y^{33}	"	Sum	2.168	-0.029	-0.053	-0.887	-1.039	-1.124	-0.015
y^{41}	D^4	AR(1)	9.775	0.784	0.515	-0.293	-0.324	-0.347	-0.006
y^{42}	"	AR(2)	8.433	0.581	0.190	-0.897	-1.017	-1.132	-0.026
y^{43}	"	Sum	7.995	0.418	0.189	-0.886	-1.036	-1.151	-0.029
y^{53}	D^5	Sum	9.553	9.306	4.951	-0.893	-0.989	-1.124	0.914
y^{63}	D^6	Sum	0.472	-0.033	-0.051	-0.887	-1.039	-1.131	-0.008
y^{74}	D^7	i.i.d.	161.0	17.87	18.40	-1.483	-1.472	-1.873	3.748
y^{75}	"	AR(1)	153.9	17.76	18.71	-7.443	-7.967	-9.404	3.525
y^{76}	"	AR(2)	157.3	19.19	20.16	-10.02	-10.97	-12.77	4.098

Table 3
Relative Serial correlation of cycle

Var.	Trend/cycle		Detrending method						
			Linear	Quadratic	Cubic	Hodrick-Prescott	Baxter-King	Christiano-Fitzgerald	Proposed k=16
y^{11}	D^1	AR(1)	-0.002	-0.004	-0.015	-0.151	-0.180	-0.176	-0.009
y^{12}	"	AR(2)	-0.001	-0.002	-0.002	-0.052	-0.068	-0.078	-0.001
y^{13}	"	Sum	-0.003	-0.005	-0.005	-0.078	-0.104	-0.118	-0.003
y^{21}	D^2	AR(1)	0.237	-0.004	-0.015	-0.150	-0.178	-0.188	-0.016
y^{22}	"	AR(2)	0.035	-0.002	-0.002	-0.052	-0.068	-0.078	0.000
y^{23}	"	Sum	0.047	-0.005	-0.005	-0.078	-0.104	-0.120	-0.005
y^{31}	D^3	AR(1)	0.257	-0.010	-0.016	-0.151	-0.180	-0.162	0.013
y^{32}	"	AR(2)	0.041	0.002	-0.002	-0.052	-0.068	-0.078	-0.002
y^{33}	"	Sum	0.070	-0.003	-0.005	-0.078	-0.104	-0.116	-0.001
y^{41}	D^4	AR(1)	0.277	0.165	0.130	-0.152	-0.172	-0.195	-0.004
y^{42}	"	AR(2)	0.057	0.020	0.007	-0.052	-0.067	-0.078	-0.001
y^{43}	"	Sum	0.095	0.022	0.010	-0.078	-0.103	-0.121	-0.003
y^{53}	D^5	Sum	0.097	0.096	0.086	-0.078	-0.093	-0.116	0.038
y^{63}	D^6	Sum	0.028	-0.004	-0.005	-0.078	-0.104	-0.117	-0.002
y^{74}	D^7	i.i.d.	0.807	0.431	0.438	-0.058	-0.055	-0.077	0.133
y^{75}	"	AR(1)	0.216	0.110	0.113	-0.102	-0.110	-0.141	0.030
y^{76}	"	AR(2)	0.077	0.042	0.043	-0.065	-0.074	-0.098	0.011

Table 4

10 Quarters Change in Trend Growth Relative to Cycle

Var.	Trend/cycle		Detrending method						
			Linear	Quadratic	Cubic	Hodrick-Prescott	Baxter-King	Christiano-Fitzgerald	Proposed k=16
y^{11}	D^1	AR(1)	0	0.000	-0.001	0.370	0.390	0.170	0.000
y^{12}	"	AR(2)	0	0.000	-0.002	0.700	0.794	0.201	0.000
y^{13}	"	Sum	0	0.000	0.000	0.610	0.710	0.247	0.000
y^{21}	D^2	AR(1)	0	-0.002	-0.003	0.373	0.402	0.205	0.016
y^{22}	"	AR(2)	0	0.001	0.000	0.702	0.800	0.204	0.010
y^{23}	"	Sum	0	0.001	0.001	0.612	0.713	0.259	0.007
y^{31}	D^3	AR(1)	0	0.0016	0.0022	0.368	0.406	0.123	0.0329
y^{32}	"	AR(2)	0	-0.0018	-0.0068	0.696	0.795	0.189	0.0051
y^{33}	"	Sum	0	-0.0010	-0.0029	0.607	0.710	0.229	0.0058
y^{41}	D^4	AR(1)	0	-0.038	-0.047	0.376	0.489	0.221	0.019
y^{42}	"	AR(2)	0	-0.012	-0.025	0.694	0.820	0.186	0.021
y^{43}	"	Sum	0	-0.011	-0.018	0.607	0.726	0.255	0.016
y^{53}	D^5	Sum	0	0.000	-0.014	0.683	0.949	0.229	0.019
y^{63}	D^6	Sum	0	0.000	0.001	0.611	0.711	0.243	0.002
y^{74}	D^7	i.i.d.	0	-0.029	-0.028	0.080	0.097	0.029	0.009
y^{75}	"	AR(1)	0	-0.027	-0.026	0.298	0.311	0.199	0.022
y^{76}	"	AR(2)	0	-0.028	-0.027	0.507	0.607	0.405	0.032

Table 5
Time Series Analysis of Artificial Data

Series: Regressor:	Δy^{21}	Δy^{22}	Δy^{23}	Δy^{31}	Δy^{32}	Δy^{33}	Δy^{63}
C	1.702	0.975	1.541	-3.710	-4.485	-6.808	0.433
	0.252	0.203	0.315	3.082	2.304	3.600	0.228
Δy_{t-1}	-0.168	0.405	0.144	-0.169	0.398	0.134	0.163
	0.070	0.065	0.070	0.070	0.065	0.070	0.070
Δy_{t-2}	-0.161			-0.164			
	0.070			0.069			
y_{t-1}	-0.023	-0.024	-0.049	-0.019	-0.021	-0.032	-0.082
	0.016	0.013	0.020	0.014	0.010	0.016	0.027
Trend	0.016	0.018	0.038	0.021	0.020	0.030	0.070
	0.014	0.011	0.017	0.011	0.009	0.013	0.022
C	1.093	0.480	0.724	1.018	0.479	0.712	0.744
	0.119	0.085	0.122	0.118	0.084	0.121	0.122
Δy_{t-1}	-0.151	0.415	0.132	-0.116	0.411	0.136	0.122
	0.070	0.064	0.070	0.070	0.064	0.070	0.070
Δy_{t-2}	-0.144			-0.106			
	0.070			0.070			

Standard Errors under estimates. Except for observations needed for initial lags, sample extends for 205 observations

Table 6
Application to Aggregate Data

	Relative Std. Dev.	Correlation with US GDP	Correlation (c_t, c_{t-1})	Correlation (c_t, c_{t-16})	Mean square change in trend growth	λ
			Hodrick-Prescott method			
U.S. GDP	1.00	1.00	0.84	-0.04	5.E-08	
Consumption	0.46	0.79	0.82	-0.16	2.E-08	
Investment	4.03	0.83	0.81	-0.12	5.E-07	
Military purchases	4.54	0.21	0.92	-0.29	2.E-06	
Nonfarm Hours	0.98	0.89	0.88	0.00	5.E-08	
Productivity	0.47	0.27	0.66	0.04	9.E-09	
Real Product Wage	0.33	0.09	0.65	-0.12	8.E-09	
Labor Share	0.46	-0.20	0.73	0.03	9.E-09	
Industrial Production	2.05	0.93	0.84	0.00	2.E-07	
Unemployment rate	45.1	-0.88	0.88	0.00	9.E-05	
Japanese GDP	0.92	0.15	0.75	-0.28	1.E-07	
			Proposed method			
U.S. GDP	1.00	1.00	0.93	-0.23	1.E-09	70235
Consumption	0.49	0.71	0.93	-0.25	1.E-09	27294
Investment	3.25	0.71	0.86	-0.23	3.E-09	78164
Military purchases	7.15	0.49	0.96	-0.36	3.E-08	64075
Nonfarm Hours	0.90	0.89	0.93	-0.22	6.E-10	80515
Productivity	0.45	0.44	0.83	-0.13	6.E-10	73231
Real Product Wage	0.44	0.49	0.91	-0.15	1.E-09	72645
Labor Share	0.42	0.05	0.86	-0.10	1.E-10	97558
Industrial Production	1.83	0.88	0.91	-0.19	4.E-09	53427
Unemployment rate	41.3	-0.91	0.94	-0.23	3.E-06	33374
Japanese GDP	1.46	0.23	0.93	-0.18	2.E-08	28291