A Method for Decomposing Time Series into Trend and Cycle Components *

Julio J. Rotemberg
Harvard Business School

October 13, 1998

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 nonparametric method for obtaining the trend ensures that short term changes in trend growth are not associated with the current level of the cycle, thus ensuring a modicum of independence between the two series. This 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. The resulting error in measuring the trend, particularly away from the boundaries of the sample, is much lower than the error from either a linear trend or a trend obtained by the Hodrick-Prescott method (which leads to a strong association between the changes in the constructed trend and the constructed cycle). Also, VAR's that involve the cyclical variables constructed by this method yield accurate representations of the behavior of the underlying cycles of several variables. Dickey-Fuller tests suggest that the artificial series have unit roots. VAR's with the series in differences, however, give a poor description of the cyclical properties of the series. I apply my method to some well known aggregate time series and show that it leads to conclusions that differ from those obtained when detrending by other methods. For example, the method suggests that real wages are quite procyclical in the U.S. and that the reduction in the growth of trend GDP in the U.S. started well before 1973.

*I wish to thank Michael Woodford for many though-provoking conversations and the Harvard Business School Division of Research for research support..

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"). It then shows that this method does a good job of separating these two components in some artificial examples where the constructed series are indeed the sum of smooth functions of time and a low order autoregressive process. Finally, it applies the method to some well known aggregate time series and discusses some of the properties of the resulting trends and cycles.

There are several reasons for filtering economic time series to extract a "cyclical" component before the statistical analysis of these series. The most important of these is that many underlying moments of economic time series, such as their unconditional mean and variance, seem to vary over time. These nonstationarities invalidate some of the properties of standard statistical techniques so that series are often transformed to render them stationary. A great many filters, including the differencing filter applied a judicious number of times, induce such stationarity. Thus, this problem, by itself is not difficult to solve.

However, the stationary time series that results from this transformation need not be identical to the true cyclical component, even assuming that such a component exists. Obtaining a good estimate of this true cyclical component is the aim of this paper. Assuming the method succeeds, it thereby also leads to a good estimate of the true trend component. The result is that one has two time series that both contain valuable (and different) information about the underlying properties of the economy.

Some interesting questions can then be asked of each. To begin with, the dating of important changes in cyclical and trend components may provide clues as to the underlying causes of these changes. One is also interested in the persistence of cyclical movements, in the relative volatility of the cyclical components of different series as well as in the comovement of these cyclical components. In particular, researchers are often interested in knowing the extent to which an unexpected change in one series is followed by cyclical movements in another. In the literature, different techniques have tended to be used for dealing with these different questions. Hodrick-Prescott (1980) filtering has tended to be used to compute the relative variability of the cyclical components of different series whereas linear trends and various degrees of differencing have been used in vector autoregressions to measure the dynamic effects of shocks of one series on another. A proper measure of the cyclical variation in a series could be used for both types of analysis, however.

The method for decomposing time series that I propose implicitly makes some assumptions about trends and cycles. I also impose these assumptions on the artificial data by which I judge the method's success. It thus seems worthwhile to start by discussing these assumptions. My first assumption, and this seems to be implicit in most definitions of trends and cycles, is that the cyclical position of a series is temporary. A high cyclical position is expected to be followed by cyclical declines that take place relatively quickly. This is closely related to the even more widely held view that the cyclical component of a time series is stationary (so that, in the long run, it is expected to equal its unconditional mean which can be normalized to zero).

Second, trend movements are smooth. By this I mean that the changes in the trend value of a series from one period to the next are almost identical to the trend changes in the preceding and the subsequent periods. Thus, changes in the rate of growth of the trend are very small. The idea that secular changes in economic can 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 of the smooth line at any date may be called the *trend* of the series at that date" (italics in original).[1] Similarly, Malinvaud (1980 p. 425) says "The *trend* is a slow variation in some specific direction which is maintained over a long period of years." The idea that

[1] 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.

trends are smooth is also broadly consistent with the way the Hodrick-Prescott (HP) trend is constructed. The HP trend is computed by minimizing the sum of two terms, the first of which is the sum of the squared changes in the rate of growth of the trend component.

Several reasons can be given for the idea that strongly persistent 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. Even if attitudes towards procreation change drastically after some event (like a war), the number of babies increases only slowly at first because only a fraction of the population is in a position to procreate at any point in time and procreation is somewhat random. As more couples find themselves in a position to have children and as those that fail at first keep trying, the population can increase dramatically even though there is no particular moment where the rate of growth changes dramatically. This is reflected in statistics of the rate of growth of the rate of growth of the U.S. population over 16. After converting this series to the quarterly frequency, it does retain a few 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.[2]

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)). Finally, the stock of knowledge that is actually used in the production of goods may well evolve smoothly even though society's knowledge is subject to breakthroughs. Many authors have emphasized (see Mansfield

[2]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.

(1968) for example) 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 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.

Third, I suppose that trend and cycle movements are independent at least in the sense that one does not expect substantial changes in trend growth in the immediate aftermath of either cyclical booms or recessions. A weakened version of this assumption actually follows from the previous one. If changes in trend growth develop only gradually, cyclical movements cannot take place at the same time as large changes in the trend. By the same token, if one expects long terms movements in hours worked to be due to the sort of demographic changes I described above, one expects cyclical movements to be due to rather different causes. And, indeed, there exist a great many macroeconomic models in which the presence of frictions leads various shocks to have temporary effects on hours worked. The combination of these shocks with slow-acting demographic shifts naturally implies that cyclical and trend movements are independent.

There is also a separate reason to be interested in situations where trend movements are at least somewhat independent of cyclical movements. Without a modicum of independence, the distinction between trend and cycle does not seem very meaningful from an economic point of view. If trend and cycle movements are of similar order of magnitude and take place at the same time, one obtains a more accurate representation of the series if the two kinds of movements are described together.[3] Moreover, theoretical explanations of cyclical movements would then have to explain some trend movements as well to be considered

[3] This is indeed what is done in the important literature that follows Hamilton (1989) and explains GDP growth by supposing that it is subject to regime shifts. Each regime is associated with a rate of growth of GDP and recessions are regimes where the associated growth is small.

attractive.

The two assumptions I make about trends, namely that they are smooth and independent of the cycle are implicit also among researchers who model the trend as a polynomial of time.[4] This approach goes back at least to Moore (1919), who recommended that a cubic trend be fitted by least squares. My nonparametric approach offers several advantages over this parametric approach, however. First, just as the freehand technique of Bowley, it is "more sensitive to changes in direction of the trend,"(Bowley, 1920, p. 137-8). Second, it avoids the implausibility of the long run forecasts that emerge from fitting polynomial time trends. Even if one properly sees this approach as fitting the polynomial trend only in sample, there is nothing particularly compelling about using time square as opposed to time to the power one half as the term one includes after one includes a linear term. It is hard to argue that one of these terms is smoother than the other. Moreover, because time ultimately changes by a large amount, one cannot say that certain polynomial terms of time must be small unless others are large.

The paper proceeds as follows. In Section 1, I present my method for detrending. This is a variant of the Hodrick-Prescott method in which it is possible for the cyclical level of a variable to be unrelated to short term changes in the trend growth of this variable. As I discuss, the method involves choosing 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 which one wants the estimated trend and cycle to be orthogonal.

Section 2 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. It is these latter processes which I treat as the cycles of the series. Section 3 begins the analysis of these data by discussing the effect of varying the parameter k on the covariation of trend and cycle at different horizons. Choosing a relatively high value of k, which I argue in Section 1 is

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

attractive on *a priori* grounds, also makes it possible to ensure that trend and cycle are close to orthogonal over a broad range of horizons.

Section 4 evaluates the success of the method with the artificial data. Except at the boundaries of the sample, the method proves vastly superior to both linear detrending and the standard Hodrick-Prescott filter at recovering the true trend and cycles. The result is that the method also produces more accurate "cyclical statistics" concerning the cyclical variability of different series as well as their correlation at different horizons. Section 5 shows that the series detrended with my method also produce quite accurate impulse responses in a vector autoregression.

That section also compares my method more generally to the approach to achieving stationarity championed by Box and Jenkins (1970) and widely adopted in practice. This approach favors the use of the first difference filter when statistical tests seem consistent with the presence of unit roots. Because this approach is more an art than a science, my attempt to analyze my artificial series with this approach and compare the results to the behavior of the true series may be somewhat idiosyncratic. Still, I show that differencing series when augmented Dickey-Fuller unit root tests fail to reject the hypothesis that such a root is present leads to quite inaccurate representations of the true series. For example, this method 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.

Section 6 finally applies my method to actual aggregate data. I show that many of the business cycle facts stressed in the real business cycle literature are in fact robust to the use of my method. On the other hand, there exist important substantive issues for which the answers obtained with my method seem quite different from those obtained by other methods. In particular, real wages seem much more procyclical when studied along the lines I suggest. Section 7 concludes.

1 The Method

I consider a time series whose observation at t is y_t where t extends from 1 to T and where this is often most usefully thought of as the logarithm of an underlying variable. I wish to decompose this series 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.

I construct the trend series d_t by minimizing

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

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

$$\sum_{t=k+v}^{T-k-v}(y_t - d_t)[(d_{t+v} - d_t) - (d_t - d_{t-v})] = 0 \tag{2}$$

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 (1980) since they minimize (1) 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 (2) - which still requires a choice of v.

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

$$\begin{aligned}[(d_{t+2} - d_{t+1}) - (d_{t+1} - d_t) &- 2[(d_{t+1} - d_t) - (d_t - d_{t-1})] + [(d_t - d_{t-1}) - (d_{t-1} - d_{t-2})] \\ &= \frac{1}{2\lambda}\Big[(y_{t+k} - d_{t+k}) + (y_{t-k} - d_{t-k})\Big] \end{aligned} \tag{3}$$

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

$$d = A^{-1}By \tag{4}$$

For given values of k and λ, the trend is thus obtained by a linear filter. This filter belongs to the Hodrick-Prescott 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 1 which gives the gain of both filters when λ is set equal to 600500, a value that turns out to be appropriate for one of the artificial series I consider in the next section. 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.

My choices of k and v are partially based on the performance of this method for the artificial data I consider in the next section. Before studying this performance, I make some general comments about the procedure which ought to help justify the choice of a strictly positive k. 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.[5]

Minimizing the expected square changes in the change of the trend makes sense if one knows *a priori* that the trend is indeed smooth. The trouble is that, if one knows only that trends are smooth, one is tempted to set the rate of change of trend changes equal to zero and let the trend be linear. One further advantage of the linear trend is that the orthogonality condition (2) is then automatically satisfied. For many aggregate time series,

[5]If ν where known for a particular horizon (as for example if it were known that the covariance of the cycle at t and the cycle at $t + k$ is zero), it might be preferable to impose this requirement and let the data determine whether the true trend and cycle are orthogonal.

the use of a linear trend poses some problems, however. These stem from the fact that the mean change in the log of many series differs in different subsamples, so that the estimated trend lines often have different coefficients for different subperiods. 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 2 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. Alternatively, one can suppose that trends are not quite as smooth as deterministic straight lines. In particular, it seems worth accepting that there exist changes in the rate of growth of the trend if, by doing so, one succeeds in making cycles shorter lived and thereby correspond more closely to what is usually meant by business cycles.

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 covariance at $t+k$ ought to be small for some positive value of k. This argument provides some rationale for making this covariance as small as possible for a given sum of squared residuals in the trend, as opposed to reducing the variance of the cycle as in the HP method.

This distinction might seem unimportant since one can check the correspondence between a model's predictions about time series moments and actual time series moments using essentially any filter as long as the filtered moments are stationary and the filter is applied to both the actual data and to the data generated by the model. Naturally, one is then testing the predictions of the model for whatever frequencies the filter allows through. In testing a model which purports to explain the business cycle, all that is important is that the filtered series capture what people usually mean by the business cycle.

Here, orthogonality in the sense of (2) plays a role, however. If the filtered series do not have this property, a part of what is usually meant by the cycle (the component of output which is temporary) is captured by the trend. This means that the variance of the measured cyclical component will not capture the variance of what people might well regard as being the true cyclical component.

One feature of the orthogonality requirement (2) that deserves comment is that it seems very weak. If the trend and the cycle are truly independent, one might expect a great many orthogonality conditions between the cycle and changes in trend to hold. One might thus wish to impose several of them and abandon the view that trend and cycle are independent if the value of λ that makes one of these orthogonality conditions hold leads to violations of others. What I do in fact is to look at several different values for v in (2) so that, I carry out an informal version of this "test".

It is important to stress, however, that a constructed trend where the trend at t depends in a constant way on observations of the actual series at $t-\ell$ cannot be independent from the constructed cycle at all leads and lags even if the true trend is independent of the true cycle. 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.

In spite of this, it seems desirable to use a value of λ such that the cyclical departure of y from trend at t is not mirrored by temporary trend movements. Thus, the main purpose of (2) is to ensure that when output is cyclically high (so that its cyclical value has recently risen and will soon fall) this does not connote that the recent increase in the trend is high relative to its expected future decrease. If this condition is violated for small values of v, the two series ought definitely to be seen as capturing the same "cyclical" phenomenon.

I have also experimented with alternative orthogonality conditions that capture some desirable forms of independence between trend and cycle. One that seems particularly attractive is

$$\sum_{t=k}^{T-k-2v} (y_t - d_t)[(d_{t+2v} - d_{t+v}) - (d_{t+v} - d_t)] = 0 \tag{5}$$

For small v, this says that the cycle at t ought not to be correlated with changes in trend growth that take place soon after t.

To get a feeling for the extent to which this, or the other orthogonality conditions are violated, I construct "regression coefficients" which give the amount by which the trend is expected to change for a given level of $-c$. In the case where one is interested in (2), the relevant coefficient β_v 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). This coefficient can be computed as

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

Similarly, the coefficient giving the expected future change in the trend growth rate is β_v^f and is computed as

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

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

2 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 cyclical series C_t is given by one of three possible series. The series C_t^1 is generated by 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. I have chose this process because its fast reversion back towards the mean suggests that C_t^1 is unambiguously cyclical rather than being closely related to a trend. 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. Finally, 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.

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^5 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 a quadratic trend which is given by

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

Rather than being a concave function of time, my third trend is convex and is given by

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

This perhaps somewhat odd choice of parameters ensures that the trend is monotone increasing while also being convex over its entire range.

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

$$D_t^4 = .87t + 17.4 \cos\left(\frac{t}{20}\right)$$

The second term in this expression has a period of about 31 years while the first term ensures that the function itself is monotone.

The four 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 they also have important effects on the long run level of the series suggests that, indeed, initial conditions are 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^5 satisfy

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

where ϵ_t^5 is drawn from a normal distribution whose standard deviation is $8.2e^{-5}$. By supposing that the third difference of the trend is stochastic, I ensure that shocks with extremely small 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^5 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 3. 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.

Finally, I also consider a stochastic trend whose second difference is an i.i.d. normal variable. In particular D_t^6 satisfies

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

The main reason for analyzing this trend (which is not as smooth as the others) is that the Hodrick-Prescott filter is optimal when the trend is given but D^6 and the cycle is given by an i.i.d. random variable. For purpose of comparison, I thus briefly consider also a fourth cycle C_t^4, which is given by the independently distributed normal random variable e_t^4. The question I study in more detail below is whether the HP trend maintains its good properties relative to my method when the trend remains equal to D^6 but the cycle is more persistent than C^4.

Given six possible trends and four possible cycles, there are 24 possible constructed y_t series. Letting i represent the index of the cycle and j represent the index of the trend, we have

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

In the next section, I provide further motivation for setting k equal to 16 by studying how the results depend on k for y_t^{32}. I focus on just one series because these results extend to the others, with the partial exception of the series whose trend is given by D^4. I return to the other series, including those based on D^4 in the subsequent section where I discuss the method's accuracy.

3 The Effect of Varying k and v

Ideally, one would like the method of detrending not to be too sensitive to the choice of v since one would like near independence of the cycle and trend growth changes at various horizons. I show in this section that, for the series I have constructed, this provides a rationale for setting k to a relatively high value. In particular, when k is set to a low value, the value of λ which ensures that β_i is zero for any given i induces values of the other β_v's that are relatively high in absolute value. By contrast, when, k is set to my preferred value of 16, the value of λ which makes β_5 equal to zero leads to low values of β_v's for all relatively low values of v.

Before studying how β_v depends on v, it is worth displaying in figure 4 the way β_5 depends on λ when y_t is given by y_t^{32} and k is set equal to 16. The relationship between β and λ

turns out to have this basic pattern for every artificial series I have detrended except for the those constructed with linear time trends. It is also found in every economic time series I have analyzed with this method.

The Figure shows that this regression coefficient is negative for low values of λ. This phenomenon, which appears in every series I have studied results from the following. When λ is small, the minimization of (1) 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). Thus the covariance between one's estimate of $-c$ and the change in the growth rate of the trend is negative.

For high values of λ, the Figure shows that this coefficient is positive and declining in λ. Indeed, if λ is made arbitrarily large the minimization problem in (1) is solved by using a deterministic trend whose slope never varies. At that point, β_5 is obviously zero. For somewhat lower values of λ, β_5 is positive. This is not true in the case where the constructed trend is linear but, as I said, it is true in all other cases.

The reason for this finding 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 (1) (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 2 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 (though not by as much as the slope of the actual trend), positive values of c are associated with reductions in the trend rate of growth so that β_5 is positive.

In between the region where β_5 is negative and the one in which it is positive, there is a point where β_5 is zero. The corresponding value of λ, in this case 600500, is the one I suggest using to estimate the actual trend. It should be clear that this technique only gives an unambiguous answer when the function that gives β_v as a function of λ crosses zero only once. The intuition I just gave seems to ensure that this single crossing does take place when the true trend has a slope that changes over time. When, instead, y_t is set equal to y_t^{31} so that the cycle is the same while the trend is linear, the function that gives β_5 as a function of λ when k is 16 is given in Figure 5. This reaches zero only when λ is arbitrarily large so that this method detects properly the linearity of this trend.

In the case of a linear trend, k and v do not appear to matter, since all the values of k and v that I have tried reach the same conclusion. Matters are different when the trend is not constant. The precise value of λ that makes β_v zero then does depend on k and v. However, for k equal to 16, β_v is not very sensitive to v near the region where λ ensures that β_5 is zero. This is illustrated in Figure 6. It shows that β_5 is zero when λ equals 600500 while β_1, β_{10} and β_{25} are all minuscule at this point. It might be appear that the λ that makes β_{25} equal to zero is quite different but, in fact the resulting filters are very similar in the sense that their gains are essentially identical.

There are two effects from lowering k to 12, both of which can be seen in Figure 7. First, the λ that makes β_5 equal to zero falls to 435000. Second, the value of β_v becomes somewhat more sensitive to v. In particular, the figure shows that β_{25} is now somewhat larger when λ is set at the value that makes β_5 equal to zero. Still, neither of these effects is large. To see this, figure 8 shows the gain from the filter one obtains when k is 16 and λ is 600500 as well as that one obtains when k is 12 and λ is 435000. The former is what one would be led to if one set k to 16 and v to 5 while the latter would follow from keeping v equal to 5 but letting k be 12. This relative insensitivity of the filter to either k or λ in this range is reassuring.[6]

Further lowering k until it equals 4 makes a much greater difference as is illustrated in

[6] For reference purposes, the Figure also shows the gain from the standard HP filter for quarterly data.

Figure 9 which shows, once again, the region of λ's where β_v is near zero. These λ's are considerably lower and whatever λ leads one of these β_v's to be zero implies that others quite large. For example, the λ that makes β_5 equal to zero (which has λ equal to 25.4) implies that β_{10} equals .64. The resulting filter is also extremely different as is illustrated in Figure 8. Because λ is so much lower, the filter that creates the trend with k equal 4 allows through much higher frequencies than the filters that make trend and cycle independent when k is either 16 or 12.

These effects of lowering k can be interpreted as follows. When k is low, the objective function (1) is reduced by having the trend rise at least slightly whenever there is a persistent short run increase in the series. As long as short term increases are somewhat persistent, this short run increase in the trend lowers the covariance of the cycle at t with the cycle at observations near t. This effect is not nearly as important when k is large because the objective function then gains much less from attempting to track individual bumps in the original series.

Because lower values of k induce the trend to get closer to the short term bumps in the series (even if only to a small extent) the high values of λ that made β_5 zero when k was 16, make β_5 positive when k is set at a lower value. With a lower k, positive bumps raise the trend and cycle together so the cycle is positively correlated with expected declines in trend growth. By lowering λ one can ensure that any particular β_v is zero for the reasons discussed above. However, the low value of λ that achieves this implies that the constructed trend follows the actual series quite closely. The change in trend growth over a particular horizon can still be orthogonal to the cycle but this is mostly an artifact rather than reflecting any true independence.

To see this, Figure 10 shows the actual value of y_t^{32}, the constructed trend when k is 16 (which has λ equal to 600500) and the constructed trend when k is 4 (which sets β_5 to zero by letting λ equal to 25.4 and which hugs the series quite closely). Even with $k = 4$, the cycle and the trend are orthogonal over 5 quarters in the way I have defined it; but they are clearly srtrongly associated in this case. By contrast, the trend obtained with k set to 16 appears in the figure to be quite distinct from the cyclical fluctuations.

I have also considered the relationship between expected future changes in the trend and the current cycle given by β_v^f. As a function of λ, this does not generally cross zero once as does β_v. Nonetheless, this measure of the relationship of the trend and the cycle also tends to be small when k is 16 and β_5 is small. In the case of y_t^{32}, setting k equal to 16 and λ equal to 600500 leads β_1^f, β_5^f, β_{10}^f and β_{25}^f to equal $1.4e^{-5}$, -$6e^{-4}$, $4e^{-3}$, and .025 respectively. By contrast, low values of k ensure that this slope is high for certain values of v even if β_5 is zero. With the same series, k set equal to 4 and λ equal to 25.4, the corresponding values of β_v^f are -.02, .28, .59 and .07.

If one reduces k until it equals zero, one is effectively choosing among Hodrick-Prescott filters with different smoothing parameters. Unfortunately, no smoothing parameter other than arbitrarily large ones makes (2) hold. When k is set to zero, the objective function (1) penalizes the variance of the cycle and thus leads the trend to rise, at least imperceptibly, whenever the cycle is high. The result is that the change in the trend is always negatively correlated with the current cycle. When the smoothing parameter is set to the usual value of 1600, the connection between the two series is quantitatively important. For y_t^{32}, β_5 is .11 and β_{16} is .42. Supposing one is dealing with logarithms, this means that when the series is 1 percent above trend, the trend itself is expected to fall by about .4 percent (relative to its previous increase) in the next four years. Forecasted future changes in trend are also quite correlated with the cycle and while β_5^f equals .09, β_{16}^f equals -.71. Because the trend is expected to slow down when the economy is in a boom, trend growth starting four years after a boom is expected to be considerably faster than trend growth over the intervening four years. Very similar coefficients both for β_v and β_v^f are obtained when the HP filter with λ equal to 1600 is applied to standard U.S. macroeconomic time series such as GDP.

This section suggest that setting k to a relatively high value ensures that the λ that makes trend and cycle close to orthogonal at one short horizon also makes them orthogonal at other relatively short horizons. I suspect this result obtains because the true trends I have considered (with the exception of D^4) have relatively small changes in their rate of growth over 16 quarters and because the cycles I have constructed have the property that their realization at t is essentially orthogonal to their realization at $t+16$. Thus, the reason for the success of setting k equal to 16 in this section may well be closely related to the *a priori* reasons I gave for doing so in Section 1.

I thus maintain k equal to 16 from now on and turn next to an evaluation of the trends and cycles produced by this method. To get a sense of the method's accuracy, I compare its trends with those one obtains from linear detrending and Hodrick-Prescott filtering in the next section. In the subsequent section, I analyze how the inferences from a vector autoregression using the constructed cycles differs from the inferences one would draw from differenced data.

4 Overall Accuracy of the Method with the Artificial Data

I start by analyzing the overall fit of the trends by presenting the mean square error of these estimates as well as the variance of these errors. Later, I turn to the perhaps more important discussion of the stochastic properties of the cycles that result from applying the method. 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 $D_t - d_t$.

Table 1a shows, for most of the artificial series, both the mean square error and the variance of the error for the trend estimates that result from the three detrending methods I consider. The mean square error is undoubtedly the best measure of overall fit. However, particularly if one is interested in studying the properties of the deviations of the cycle from its mean value, the mean error does not matter while the variance of the error does.

The first three rows of results for each detrending method give information about what happens when the true trend is linear. My method then essentially fits a linear trend which, like the trend that results from running a regression on a linear trend, is very close to the true trend.[7] The closeness of the fit is not surprising since, in either case, one is essentially

[7] The use of my formulae is awkward, however, because the matrix A in (4) ceases to be invertible if λ is

fitting a two parameter model (the position and the slope) of the true trend.

The next three rows give the results from detrending variables whose trend is quadratic (given by D^2) while the following three rows give the results from detrending variables whose trend is given by D^3. It is apparent from the table that the linear trend is very unsatisfactory in these cases and that it leads to large errors which are essentially independent of the stochastic process for the cycle.

Further insight into the fit of both the linear trend and the trend I propose can be obtained from Figure 11 displays the time series for these errors in the case that the trend is given by D^2. The three lines that trace out a nearly identical parabolas are the errors from the linear trend. The errors from applying my method are considerably smaller, particularly when the cycle is given by C_t^1, though also in the two other cases. Not surprisingly, most of the error in the trends considered here occurs near the edges of the sample. This occurs because the trend is both very smooth (so that its changes in the rate of change from one period to the next are small) and subject to large changes in both the level and the rate of change over longer periods of time. Thus, there is little information about the true trend contained in the data at the beginning and the end of the sample. It is important to note that the relatively poor quality of the trend estimates at the boundaries hinges mainly on λ and is not importantly affected by k. If k is set to zero and λ is kept at 600500, the trend is very similar (as is suggested by the similarity of the filters displayed in Figure 8) and results in the same errors at the beginning and the end of the sample.

The effect of these difficulties at the boundaries is that the mean square error falls to between one third and one quarter of its full-sample value if one neglects the first and the last 16 observations. This is what I do in the last two columns of Table 1. This dramatic reduction in the mean square error obtains even though the mean of the estimated trend tends to be further away from the mean of the true trend. Nonetheless, the variance of the error falls so dramatically that the mean squared error falls a great deal.

made arbitrarily large. I have thus used the largest possible value of λ which did not yield an error message. Given that the correlation between trend changes and cycles is negative even for these highest possible values of λ, my method really ought to lead to linear detrending in this case.

One further aspect of note is the dependence of the fit of the trend on the nature of the true cycle. With linear detrending, the trend fits so badly that the nature of the cyclical fluctuations has little effect on its mean square error. For the trend I propose, however, the mean square error is bigger for C^2 than for C^1 and larger still for C^3. To some extent this may result from the differences in the variance of these cyclical series; the variances of C^1, C^2 and C^3 equal 2.15, 7.62 and 9.96 respectively. In addition, the ratio of the mean square error for the intermediate sample to the variance of the cyclical component is somewhat lower for C^1 than for the other series so that other elements of the stochastic process for the cycle matter as well.

For the other trends that I have considered, the fit of the trend I propose also deteriorates near the boundaries. When the true trend is given by D^3, the qualitative aspects of the performance of both the linear trend and the trend I propose are quite similar to their performance when the true trend is D^2.

The sample paths of the errors that result from these detrending methods when the cycle is given by C^3 while the trends are given by D^4 and D^5 are given in Figure 12. The main reason to be interested in D^5 is that, in some sense at least, this trend is purely stochastic. In other words, the variance of its long run forecast is arbitrarily large. Nonetheless, initial conditions matter a great deal and, in this sense, much of what happens to the trend over the sample that I analyze is independent of the shocks over this sample. Nonetheless, the shocks do matter for, otherwise, the trend would be linear given the initial conditions I have chosen. The departures from linearity that result from the stochastic shocks in the sample are responsible for the improved fit (over a linear trend) of the trend I propose.

Where my method runs into difficulty is with the growth cycles of D^4. As Table 1 shows, this leads to a much larger mean squared error. To some extent, this is due to difficulties at the boundaries that simply extent further inwards. Thus, the mean square error falls further to 2.4 if only the period from the 21$^{\text{st}}$ to the 157$^{\text{th}}$ observation is considered. But, in addition, this is a series in which β_v is quite sensitive to v even when k equals 16. Thus, when λ is set to 10950 so that β_5 is zero, β_{10} is .017 and β_{25} is .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.[8]

I now turn to a comparison with the Hodrick-Prescott trend based on the standard smoothing parameter of 1600. Some of the characteristics of this trend are given in the bottom panel of Table 1. This shows that, at least for the series I have considered, the accuracy of this trend is essentially independent of the true trend while it depends considerably on the nature of the true cycle. The mean square error for all the HP trends for series whose cycle is C^3 hover around 1.98 for the full sample and 2.18 when one neglects the first and last 16 observations. One reason the HP filter appears to perform relatively well near the boundaries is that I set the cycle in the periods before the first observation equal to zero. This implies that keeping the variance of the cycle low at the beginning of the sample helps reconstruct the true trend. The evaluation of this method would presumably be more accurate if, instead, I drew the initial observations of the cycle from the cycle's unconditional distribution for adjacent observations. Rather than doing this, I evaluate the HP trends neglecting the observations at the boundaries as well.[9].

If one focuses on this interior region, the HP trend is less accurate as a measure of the trend for the series in Table 1a except in the case of D^4.[10] For the HP trend, the mean

[8]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 (4). 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.

[9]A related reason for doing so is that the HP trend also tends to lose accuracy near the boundary, just as the trend I propose. This is stressed in Baxter and King (1995)

[10]These inaccuracies are closely related to the distortions introduced by the filter reported in King and

error is close to zero so that the mean square error is nearly identical to the variance of the error. The result is that, while the HP trend has higher values of the mean square error, it has particularly high values of the variance of the error relative to the trend obtained with k equal 16.

Table 1b compares the performance of the two methods more directly because it also looks at a series for which the optimal detrending method is the HP filter. In particular y^{46} is the sum of a trend whose second difference is a mean zero i.i.d. random variable (e_t^6) and a cycle, which is an independent i.i.d. random variable (e_t^4) whose standard deviation is about forty times larger than the variance of e_t^6. As shown by Hodrick and Prescott (1980) the resulting optimal detrending method is the HP filter with smoothing parameter set equal to 1600. Indeed, the HP filter has much lower MSE for the trend in this series.

The other series considered in this table have the same trend D_t^5. The cycle for y^{56} 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 in y_t^{66} follows the same AR(2) perocess 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 e_t^4. The table shows that for these two cycles my method has a substantially smaller MSE in the intermediate range. Thus the use of a cycle that is more persistent (and more plausible) than an i.i.d. series implies that the HP method no longer decomposes the series very well. My method 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^{46} is set at 15637, while those for y^{56} and y^{66} 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.

Further insight into the behavior of the HP trend can be obtained from Figure 13. This gives the sample path of the error of the HP trend for the three cycle series assuming that the trend is given by D^2 (the sample paths for the other trends considered in Table 1a

---

Rebelo (1993).

are essentially identical, as I discussed above). These errors simply reflect - and are highly correlated with - the cycles themselves. One way of thinking about these errors is to view the HP trend in the way emphasized by Baxter and King (1995), namely as a series that results from the application of a low pass filter. One can then think of its inaccuracy as a measure of the trend as coming about because, when applied to these series, this filter lets through too many frequencies. In other words, the boundary frequency at which it cuts off its input is too high. The result is that cyclical movements are reflected in the trend.

To this it might be objected that Baxter and King (1995) show that the HP filter is fairly close to a filter that lets through only frequencies that correspond to periods longer than 32 quarters. If actual business cycles generally last only around 16 quarters, such frequencies might well be viewed as noncyclical. The problem is that a series such as C^1, which reverts to the mean extremely fast (so that it is expected to equal 1 percent of its original value after 16 quarters) still has a lot of power at frequencies whose periods are longer than 32 quarters. Indeed, 21 percent of its variance is accounted for by power at these frequencies.[11]

Table 2 gives a further indication of this by showing the simple correlation between the error in measuring the trend $d_t - D_t$, or $C_t - c_t$, and the actual cycle C_t. Sampling error keeps this correlation from being zero even when a simple linear trend is fitted to data that was generated by a linear trend. What the table shows is that the correlations that result from both linear detrending and from my method are roughly of the same order of magnitude (except when my method with k equal to 16 is applied to data whose trend is given by D^4). The table also shows, however, that this correlation is substantially greater in the case of the HP trend, where it reaches the value of .78 when the cycle is given by C^2.

There are a number of other ways in which one can document the extra accuracy of the cyclical series obtained by setting k equal to 16. For example, the autocorrelation function

[11] For a first order autoregressive process with first order serial correlation parameter equal to ρ, the variance accounted for by frequencies whose absolute value is below $z\pi$ equals

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

Setting z equal to 1/32 and rho equal to .75, the result obtains.

that results is much more similar to the autocorrelation function of the original C variables. Instead, the linear trend leads to measured cycles whose autocorrelations die down much more slowly (they look more like the series has a unit root). Finally, the HP trend leads to autocorrelations that die down more rapidly.

This is not, by itself, an argument for not using the HP filter in the evaluation of models as long as both the actual and the model-generated data are filtered. This is indeed the practice in evaluating real business cycle models (see for example the survey of King and Rebelo 1998). What is more problematic is to regard the resulting test as a test of the cyclical accuracy of the model, given that one is ignoring some fluctuations that are highly correlated with the fluctuations one treats as cyclical. A more appropriate test would thus analyze whether the model also leads to these correlated trend movements. Alternatively, it would focus on explaining cyclical movements that are close to orthogonal to trend fluctuations. This raises the question the question of whether the statistics that are usually computed with HP filtered data remain unaltered when the constructed business cycle is orthogonal to the trend in the way I have suggested.

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, when computed with the HP filter or other detrending methods, these statistics give accurate measures of the relevant population moments. Table 3 provides some evidence on this. The first three columns give the ratio of the computed standard deviation to the actual standard deviation of the relevant cyclical series. Because the HP detrended series is essentially independent of the trend, I provide only one measure of this ratio for the HP filter. By contrast, the computed standard deviation of the cycle does depend on the underlying trend, particularly when the series are linearly detrended though also, to some extent, when they are detrended by setting k equal to 16. In the case of these detrending methods, I thus use a separate line for series based on different trends.

The first row shows that, consistent with my earlier discussion and with the results of King and Rebelo (1993), the HP filtered data understate the standard deviation of the underlying cyclical series. Moreover, as suggested by King and Rebelo (1993), the extent of this understatement depends on the nature of the underlying cyclical series itself. If the underlying cycle is given by C^1 (which returns relatively quickly towards its mean) the understatement is smaller than in the case of C^2 which is more persistent. This makes sense given the low-pass nature of the HP filter. What this means, however, is that ratios of standard deviations based on HP filtered series can be somewhat misleading as tests of the model's accuracy.

Suppose, for example, that the model's predicted standard deviation for some series coincides with the actual one. If the stochastic process of the series predicted by the model is different from the stochastic process of the actual data, then actual and model generated "cycles" can be correlated with quite different "trend" variations over short time periods. The result is that when these correlated short term "trend" variations are included in the "cycle", the variabilities of the two cycles can be different.

Similarly, if two series measuring different economic variables have different stochastic processes, the HP filtered versions of these series are correlated with different short term trend variations. Thus, the ratio of their standard deviations might be different if these correlated trend variations were included in the measured cycle. It would seem more meaningful to test the model as a model of the business cycle by comparing either the variability of forecasted changes over some "cyclical" horizon predicted by the model to the actual ones (as in Rotemberg and Woodford (1996)) or, by comparing the variability of cyclical changes that are uncorrelated with trend changes over this horizon, as is the case of the changes captured when k is set to 16.

The last two columns focus on correlations and obtain quite similar results. Both the HP method and the method I propose correctly detect that series containing C^1 are uncorrelated with series containing C^2. I thus focus on the correlation of series that contain C^3 (which is itself the sum of C^1 and C^2) with series that contain one of C^3's components. The actual correlation between these series is given at the bottom of the table. The correlation between an HP filtered series containing C^1 and an HP filtered series containing C^3 overstates the correlation between the two cyclical series. By contrast, the correlation between an HP filtered series containing C^2 and an HP filtered series containing C^3 understates the correlation between the two cyclical series. The reason is, once again, that the HP filter treats a higher fraction of the variations due to C^1 as cyclical than it does of the variations due to C^2. Thus, the HP filtered version of series containing C^3 contains a disproportionate amount of variations due to C^1. My method, once again, obtains substantially more accurate measures of these underlying correlations.

5 Time Series Analysis of the Artificial Data

The analysis so far has focused on comparing the method of detrending I propose to other methods of detrending. In this section, I deal with a much more widely used set of methods for handling nonstationarities in aggregate data. For lack of a better name, I will call this the time series approach. A perhaps extreme view of detrending from a leading proponent of this approach is "It is common practice, particularly in economics, to detrend a series before subjecting it to analysis. The best advice to anyone contemplating such a course of action is not to do it."' (Harvey, 1989, p. 289). Rather, as suggested by Box and Jenkins (1970), nonstationarities are dealt with in this approach by suitable differencing.

Assuming it does lead to the correct characterization of the underlying process followed by the series, this approach is superior in many respects to the kind of detrending carried out here. First, it appears to avoid the need to make assumptions about the connection between trend and cycles; the data alone appears to say how secular and short term changes are connected. Thus, the independence of these changes can be tested, rather than being assumed.

Second, as argued by Diebold (1998) it is extremely attractive 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.

Third, and this is a key insight of Box and Jenkins (1970), differencing allows for a variety of trends, including some very smooth ones. For example, the series containing the quadratic trend D^2 can be represented as an ARIMA process that is I(2) while the series containing the trend D^5 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).

However when the trends are smooth as in these examples, the parameter estimates obtained by the time series approach do not have standard distributions. While it is true that series with trends that are polynomials of time can be described as ARIMA processes, the MA polynomials of the resulting processes are noninvertible if the series also contains a stationary component. In particular, the resulting MA polynomial contains as many unit roots as the degree to which one must difference the series to make it stationary. The case in which the trend is linear so that the resulting MA process has a single unit root, has been studied extensively. Plosser and Schwert (1977) already showed that, in this case, linear detrending leads to more accurate estimates than those one obtains by first differencing the series. The distribution of the maximum likelihood estimate of the MA coefficient that results from differencing this process is derived in Davis and Dunsmuir (1996). 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.

These problems pale by comparison with those I encountered when I used my commercially available software (which does not use maximum likelihood) to fit the series with quadratic trends by differencing the series twice. 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.[12] 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^5.

[12] In the case of y^{12}, for example, this is an ARIMA(1,2,2).

These issues are somewhat removed from actual practice since applied macroeconomists rarely fit MA components. This raises the question of what a typical time series practitioner would conclude when confronted with data that included a smooth trend and, in particular, what he or she would conclude about my artificial data. This question seems difficult to answer, particularly as it is likely that different investigators would do different things. The approach I will follow is based on Dickey and Fuller (1979) and Box and Jenkins (1970).

For each artificial series except that based on D^4 which I ignore in this section, I first carry out augmented Dickey-Fuller tests for the presence of a unit root. I carry out these tests including a linear trend and the result is that when the trend is given by D^1, the augmented Dickey-Fuller test correctly rejects the hypothesis that a unit root is present.[13] The results of running these augmented Dickey-Fuller regressions when the true trends are given by D^2, D^3 and D^5 are reported at the top of Table 4. It is apparent from this Table that the augmented Dickey-Fuller test leads one to accept the presence of a unit root.[14]

The bottom of Table 4 gives my estimates for stochastic processes followed by the first difference of these series. The approach to obtaining these is based loosely on Box-Jenkins (1970). The portmanteau Q statistic for misspecification does not detect any abnormal autocorrelations and this seems like a further reason to stop here. 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.

Shocks to the series would then be deemed to have permanent effects even though, as

[13] I also carried out tests without including a linear trend and obtained similar results.

[14] 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.

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^{12} 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. [15]

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 exactly the same geometric way. To conserve on space, I thus report only the response of C^3 to the two shocks in Figure 14.

I fit the same VAR to the series c_t^1 and c_t^3 that are obtained from y_t^{12} 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^{12} 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^{12} first, I ensure once again that there is no response of y_t^{12} to the shocks of y_t^{33}. Figure 14 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 response of the level of this series is obtained by integrating forward the estimated responses

[15] To both 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.

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.

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 crudely analyzed by standard time series methods.

More sophisticated methods ought to do better. However, it should be clear that application of the differencing filter - a band-pass filter that emphasizes high frequencies - is not really the same as letting "the data speak for itself" about the connection between cycle and trend. Like the method I propose, this method too seems more appropriate in some circumstances than in others.

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 by arguing that, in certain circumstances, it uncovers the properties of the underlying data when other approaches do not. This naturally leads to the question of whether the method provides a different characterization of actual business cycles than do alternative methods. This is a rather large topic so the aim of this section must necessarily be modest. I thus focus only on a few key aggregate variables and study just their cyclical variability relative to the variability of detrended U.S. GDP and their correlation with detrended GDP. This allows me to study the extent to which the "cyclical facts" that have been established by HP filtering the data remain valid when looked at through my method.

I analyze quarterly data on the log of U.S. GDP and of two of its components, namely consumer expenditure on nondurables and services, and investment, where this is defined to include consumer durables purchases as well as business fixed investment. To compare total output to a more cyclical component, I also study the log of (quarterly averages of) industrial production. I use total hours in the nonfarm sector worked as well as the log of the ratio of GDP to total hours to compare the behavior of output to that of the labor input. To analyze the behavior of the payments to labor, I compute statistics both for the log of ratio of total nominal labor compensation to nominal GDP (a measure of the labor share) and for the real wage that is obtained by dividing this labor share by my measure of aggregate labor productivity (this is obviously equivalent to the ratio of total compensation deflated by the GDP deflator over my measure of total hours). These data are all quarterly from 1947:1 to 1998:1.

I also consider two additional variables whose observations cover a shorter time period. The first of these is the U.S. unemployment rate which is available from 1948:1. I include this because 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 unemployment becomes more closely associated with other measures of the business cycle once it is detrended using either the HP filter or my method.

The last variable I consider is the log of real GDP in Japan which is available from 1957:1 to 1997:2 in the IMF International Financial Statistics. My reason for including this variable is that Taylor (1989) has argued that Japanese GDP is less cyclically variable than U.S. GDP (and has provided a model to explain this) while, at the same time, 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).

To detrend these variables I set k equal to 16 and search for values of λ such that β_5 is zero. Interestingly, the resulting values of λ all cluster within a factor of four of one another. This parameter equals 70,000 for U.S. GDP and the value for the other series are given in table 5. The fact that hours, GDP and productivity all involve such similar values of λ implies that the difference between detrended log output and detrended log hours is essentially identical to detrended log productivity.

Before discussing the other entries in the table, it is worth displaying U.S. GDP together with the trend I compute. I do this in Figure 15, This Figure shows that, as might be expected, the trend I compute grows more slowly after the early 1970's than before. Even ignoring the observations near the boundaries of the sample which are not reliable, several things are interesting about this trend, and they are particularly visible in the second panel which shows the rate of growth of this trend. The first is that the rate of growth of trend GNP actually rose in the early part of the sample. Second, as suggested by the work of Perron (1989) the early 1970's, and the period around 1973 in particular, were a moment where the rate of growth experienced an unusually sharp reduction. 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. If these results hold up, they would cast doubt on the view that the secular decline in output growth of the U.S. was caused by the oil price shock of 1973.

One remarkable aspect of Table 5 is that the business cycle facts emphasized in the real business cycle literature are robust to computing the trend by setting k equal to 16.[16] Indeed, except for the comparison between the U.S. and Japan, the ratios of standard deviations are very similar when computed with the two methods. Interestingly, Japanese cyclical GDP does prove to be more variable when computed using my method whereas it is less variable using the HP filter. The reason, not surprisingly is that changes in the HP trend are more strongly associated with the HP cycle in Japan than in the US.[17] Another contrast is that the series for the business cycle computed for the two countries with my method are more closely associated with one another.

Both methods imply that the correlations between the cyclical components of U.S. GDP and total cyclical U.S. GDP are very high. Cyclical unemployment is slightly more correlated with cyclical output with my method (thereby rationalizing the use of unemployment statistics as indicators of the business cycle). But even the HP method leads to much higher correlations between cyclical unemployment and output than linear detrending (the correlation between these two linearly detrended series from 1948:1 to 1998:1 is only -.22).

Perhaps the most important difference between the results from applying my method and those from the HP filter is that real wages are much more procyclical when the trend is removed with my method. That the extent to which real wages are deemed to be procyclical depends on the method of detrending should not be too surprising since the secular behavior of real wages seems to differ from that of GDP. Both exhibit lower average rates of growth after 1973 but the decline in average wage growth (and in labor productivity growth) are higher. Thus, one cannot get a reliable estimate of the cyclical association between these series until one removes these quite different trends from them.

[16] While this does not have a big effect on the results, these statistics are computed over the sample 1954:1-1994:1 to reduce the effect of the observations near the boundary of the sample.

[17] A regression of the difference between the rate of growth of the HP trend over the next two years minus the rate of growth in the past two years on the HP cycle has a coefficient of -.65 in Japan and of -.43 in the US.

7 Conclusions

In this paper I have provided a method for detrending time series by separating them into a smooth component and a temporary stochastic component. The method appears to do a good job of reconstructing such components when it is applied to some artificial time series which actually contain these components. Still, I have not been able to demonstrate its general properties even in this case and that is left for further research. It is also worth exploring whether modifications of the way the method treats initial conditions would lead to improved measures of the true trend near the beginning and the end of the sample. It is important to realize, however, that if the true trend changes smoothly, one cannot do a good job of detecting its level until one has seen quite a few observations beyond the current period. In other words, the fact that the trend changes vary little from one period to next even when these changes have profound long run consequences (while the cycle's period-to-period changes are correspondingly much greater) makes detection of trend changes near the end of the sample very difficult.

Another important open question, and one I have already alluded to, is how one ought to test the hypothesis that trends are smooth against the hypothesis that time series contain a single unit root (as is often deemed appropriate for aggregate time series). A start in this research project would be to extend the analysis of the Dickey and Fuller (1979) and Phillips and Perron (1988) to the case where the series is previously detrended through my method (or where the resulting trend is included as an additional right hand regressor). This would at least allow one to test the null of a unit root against the alternative suggested here. Moreover, it would then be possible to analyze the power of this test by studying the extent to which it rejects this null in examples such as the ones presented here.

8 References

Baxter, Marianne and Robert G. King, "Measuring Business cycles: Approximate Band-Pass Filters for Economic Time series," *NBER Working Paper 5022*, February 1995.

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.

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.

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, "Pot-War Business Cycles: An Empirical Investigation,", Working paper, Carnegie Mellon University, 1980.

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.

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.

Phillips, Peter C. B. and Pierre Perron, "Testing for a Unit Root in Time series Regression," *Biometrika*, 75, 1988, 335-46.

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, 199-224.

Rotemberg, Julio J. and Michael Woodford, "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.

Table 1a
Errors in Computed Trends

Variable	Full Sample MSE	Full Sample Var.	Intermediate Range MSE	Intermediate Range Var.
Trend with k=16				
y11	0.41	0.06	0.39	0.04
y21	0.26	0.00	0.26	0.00
y31	0.21	0.02	0.20	0.01
y12	0.61	0.57	0.13	0.12
y22	2.46	2.30	0.80	0.78
y32	2.95	2.71	1.01	1.00
y13	0.71	0.71	0.25	0.21
y23	0.37	0.37	0.26	0.19
y33	0.57	0.55	0.47	0.35
y34	14.09	14.08	4.37	4.22
y35	0.48	0.45	0.23	0.15
Linear trend				
y11	0.01	0.00	0.01	0.00
y21	0.22	0.19	0.17	0.13
y31	0.30	0.22	0.23	0.16
y12	22.93	22.92	13.71	11.64
y22	23.14	23.11	13.59	11.77
y32	23.22	23.14	13.37	11.79
y13	31.30	31.29	19.66	16.14
y23	31.51	31.48	20.35	16.48
y33	31.59	31.51	20.77	16.52
y34	148.72	148.64	158.11	157.81
y35	4.76	4.68	3.19	2.29
HP trend				
y11	0.40	0.39	0.35	0.35
y21	1.99	1.95	2.13	2.10
y31	1.98	1.90	2.18	2.15
y12	0.41	0.40	0.35	0.35
y22	1.98	1.94	2.13	2.10
y32	1.98	1.90	2.18	2.14
y13	0.40	0.39	0.35	0.35
y23	2.00	1.96	2.14	2.10
y33	1.98	1.90	2.19	2.15
y34	2.07	1.99	2.19	2.14
y35	1.98	1.90	2.18	2.15

Note: The full sample contains 205 observations, while the intermediate range neglects the first and last 16 observations.

Table 1b
Errors in Computed Trends

Variable	Intermediate Range MSE	Var.
Trend with k=16		
y^{46}	246.17	246.17
y^{56}	294.59	293.32
y^{66}	329.41	325.71
HP trend		
y^{46}	56.33	55.42
y^{56}	389.89	377.32
y^{66}	488.91	465.96

Table 2
Correlations between trend and cycle

Variable	k=16	linear	HP
y^{11}	-0.11	0.11	0.65
y^{21}	0.10	0.10	0.78
y^{31}	0.13	0.13	0.75
y^{12}	0.21	0.04	0.65
y^{21}	0.16	0.13	0.78
y^{31}	0.17	0.14	0.75
y^{13}	0.12	-0.01	0.65
y^{23}	0.13	-0.11	0.78
y^{33}	0.11	-0.10	0.74
y^{34}	-0.20	0.04	0.75
y^{35}	0.08	-0.11	0.75

Table 3

	Variance Ratios			Correlations between	
	C^1	C^2	C^3	C^3 and C^1	C^3 and C^2
HP	0.64	0.47	0.54	0.61	0.88
k=16					
D^1	1.00	1.00	0.99	0.52	0.90
D^2	0.96	1.00	0.99	0.53	0.90
D^3	1.02	0.98	0.99	0.54	0.90
D^4			1.63		
D^5			1.00		
linear					
D^1	1.00	0.99	0.98	0.51	0.90
D^2	6.21	2.13	1.78	0.77	0.94
D^3	8.57	3.35	2.73	0.86	0.96
D^4			15.03		
D^5			1.30		
Actual				0.52	0.90

Table 4
Time Series Analysis of Artificial Data

Series:	y^{12}	y^{22}	y^{32}	y^{13}	y^{23}	y^{33}	y^{35}
Regressor:							
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 5

	λ	Standard Deviation as ratio to S.D. of GDP		Correlation with GDP	
		k=16	HP	k=16	HP
Consumption	27000	0.52	0.48	0.81	0.81
Investment	78000	3.14	3.88	0.83	0.88
Ind. Prod.	53300	1.77	2.04	0.92	0.92
Total hours	80500	0.90	0.96	0.91	0.88
Productivity	73200	0.42	0.48	0.44	0.33
Labor share	97500	0.43	0.47	-0.07	-0.27
Real Wage	72700	0.35	0.31	0.42	0.09
Unemployment	33500	0.41	0.45	-0.92	-0.88
Japanese GDP	28600	1.19	0.88	0.28	0.16

Figure 1: Gains for Hodrick-Prescott and k=16 with λ=600500

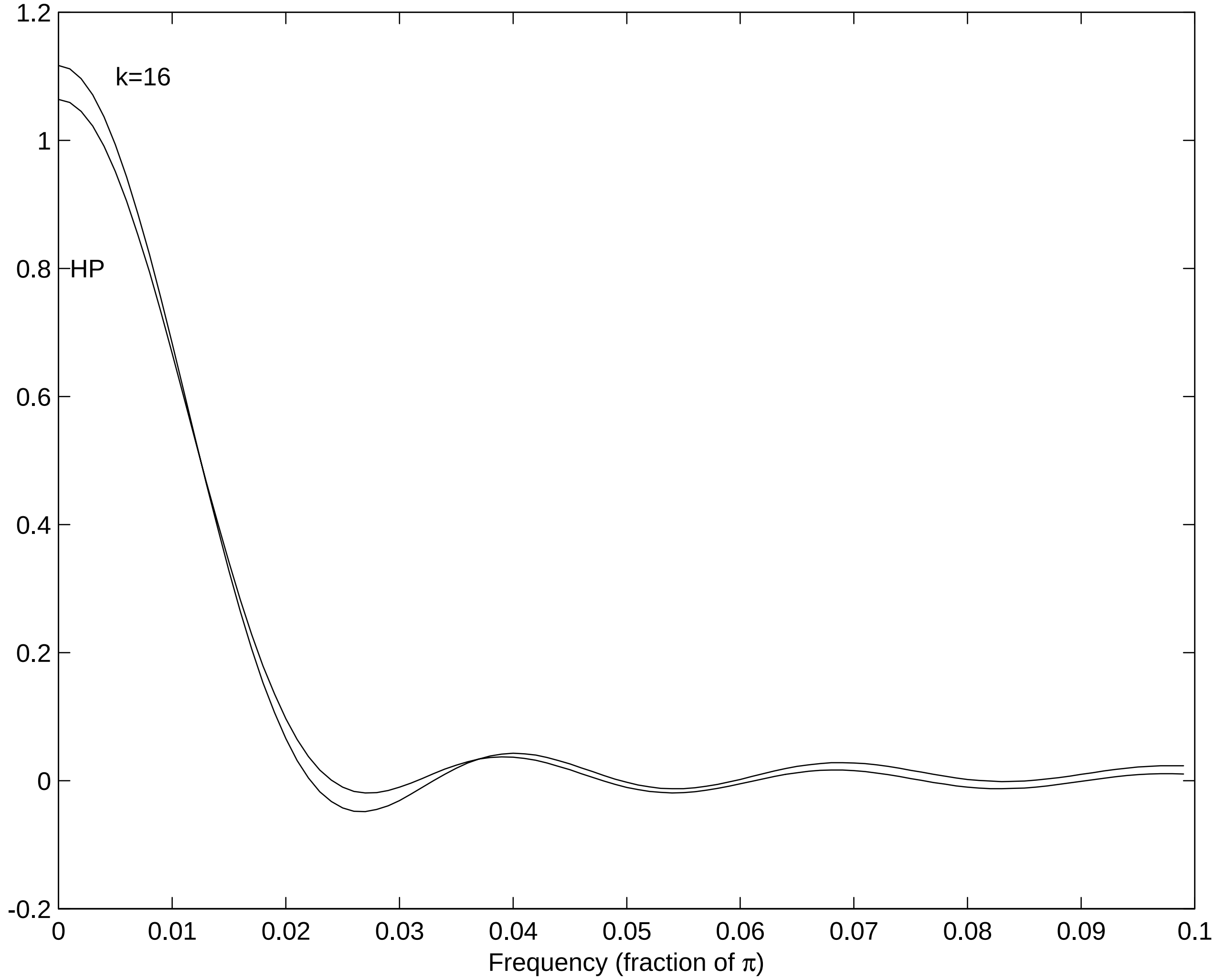

Figure 2: Log of U.S. GDP and its linear trend

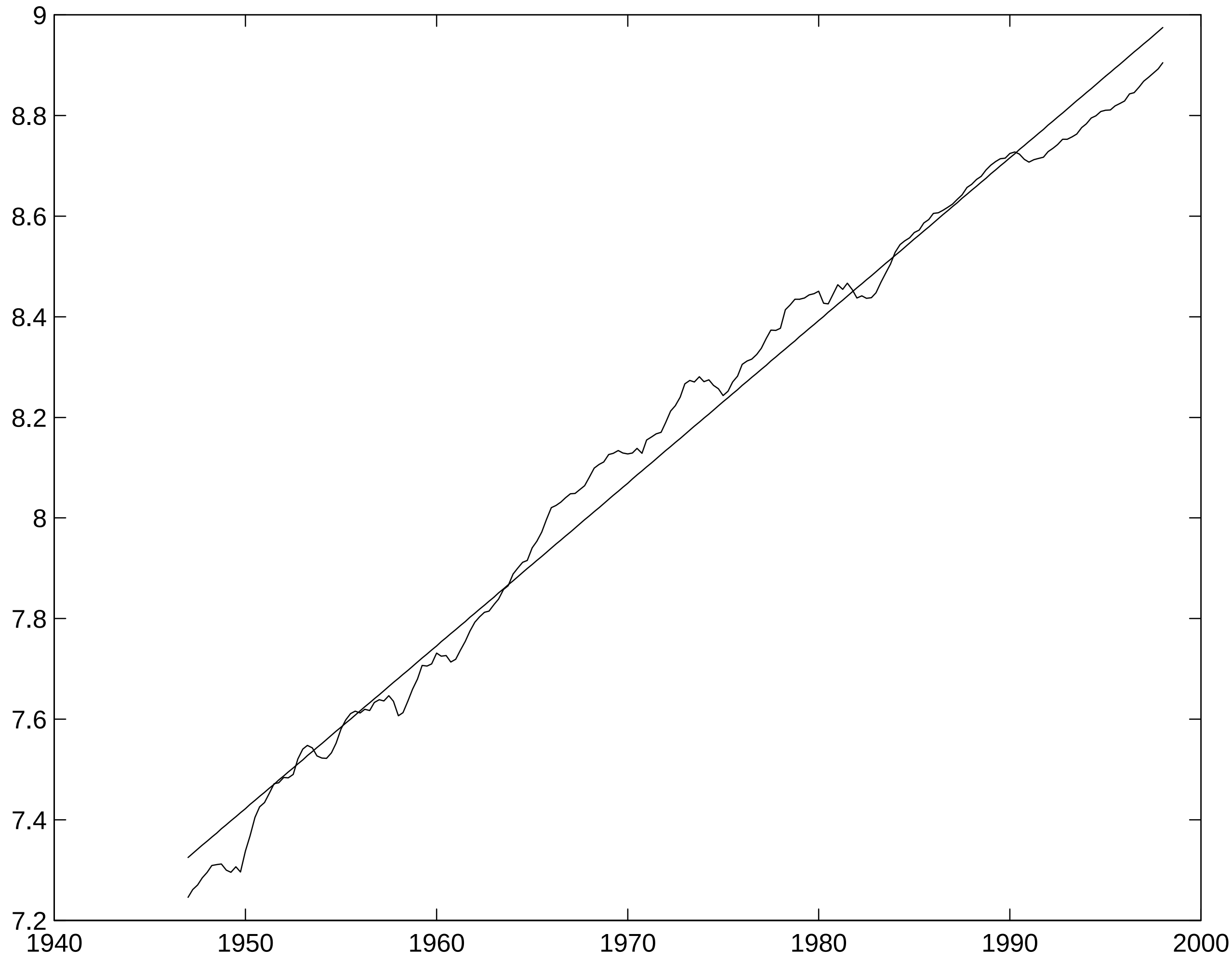

Figure 3: Artificial Trend D^5

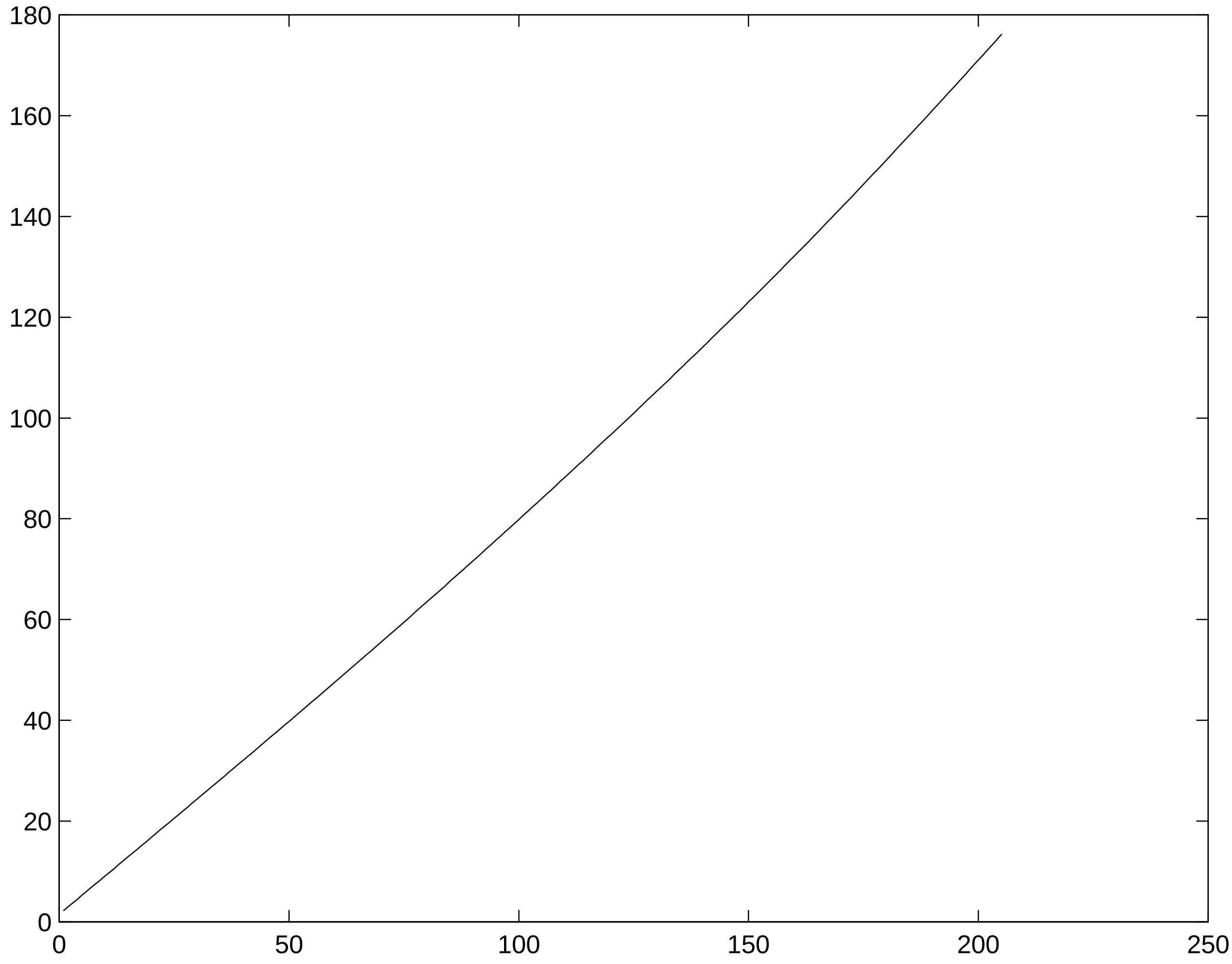

Figure 4: Regression coefficient of change in trend on cycle, y^{32}, k=16

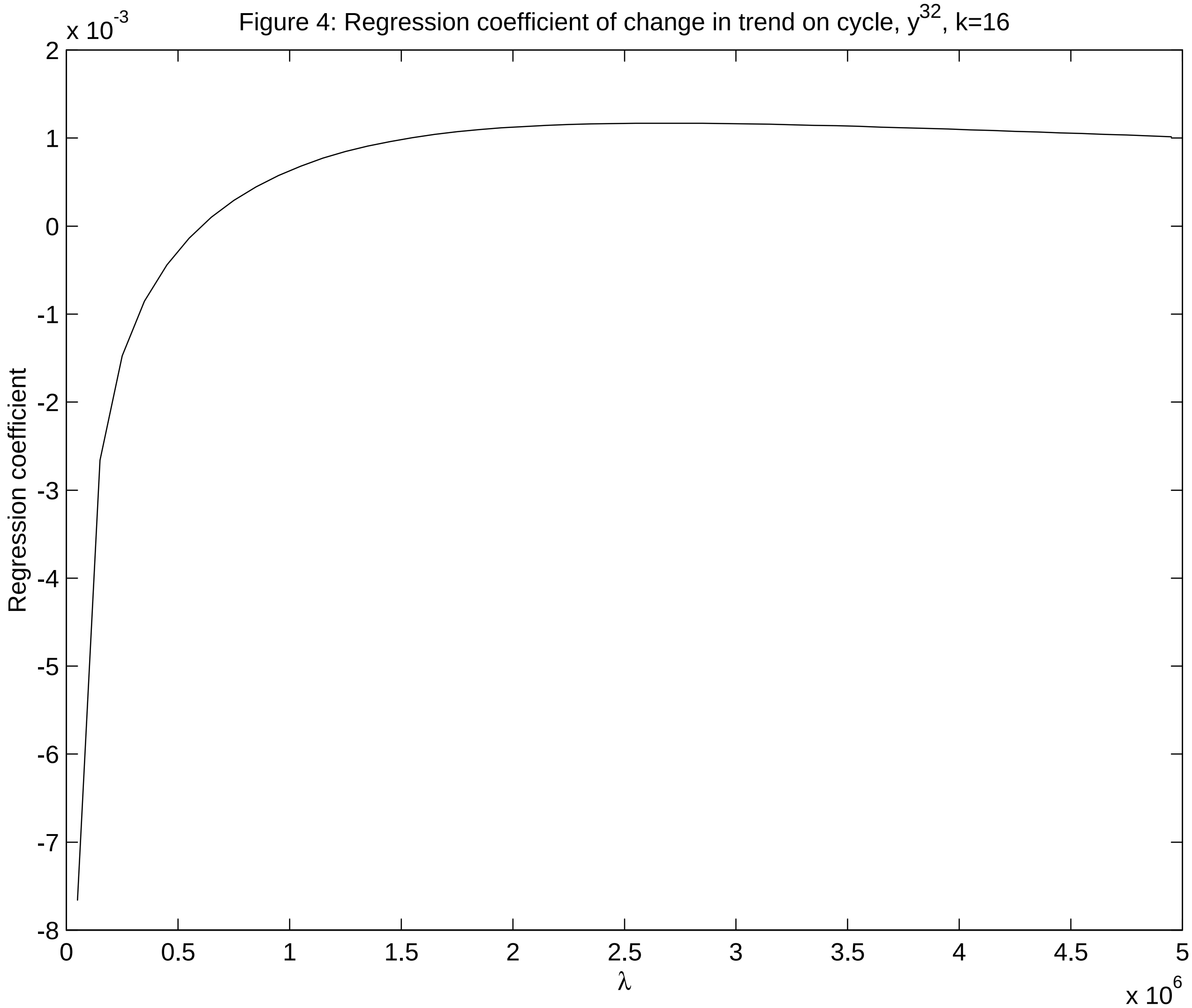

Figure 5: Regression coefficient of change in trend on cycle, y^{31}, k=16

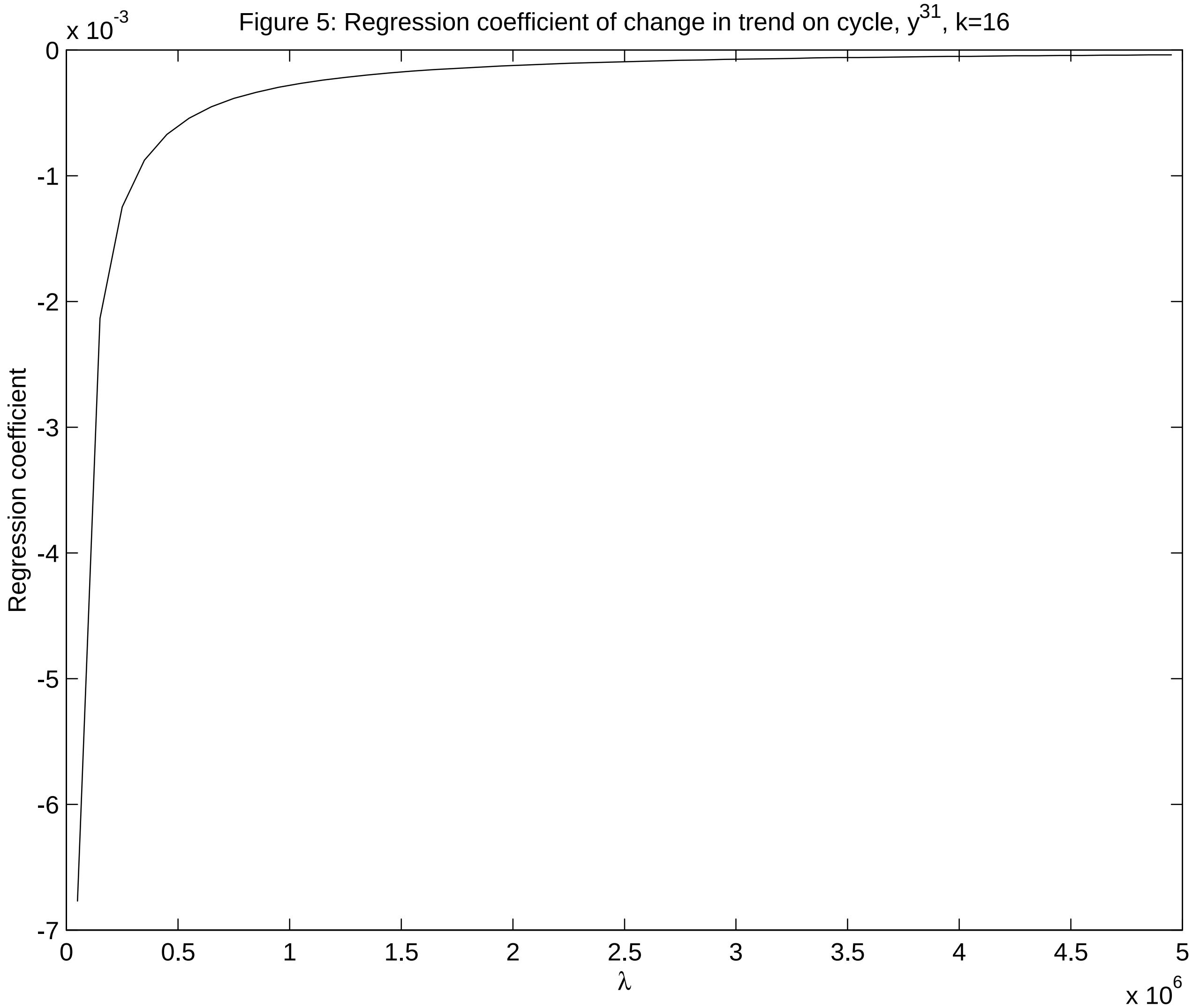

Figure 6: Regression coefficients of change in trend on cycle, k=16

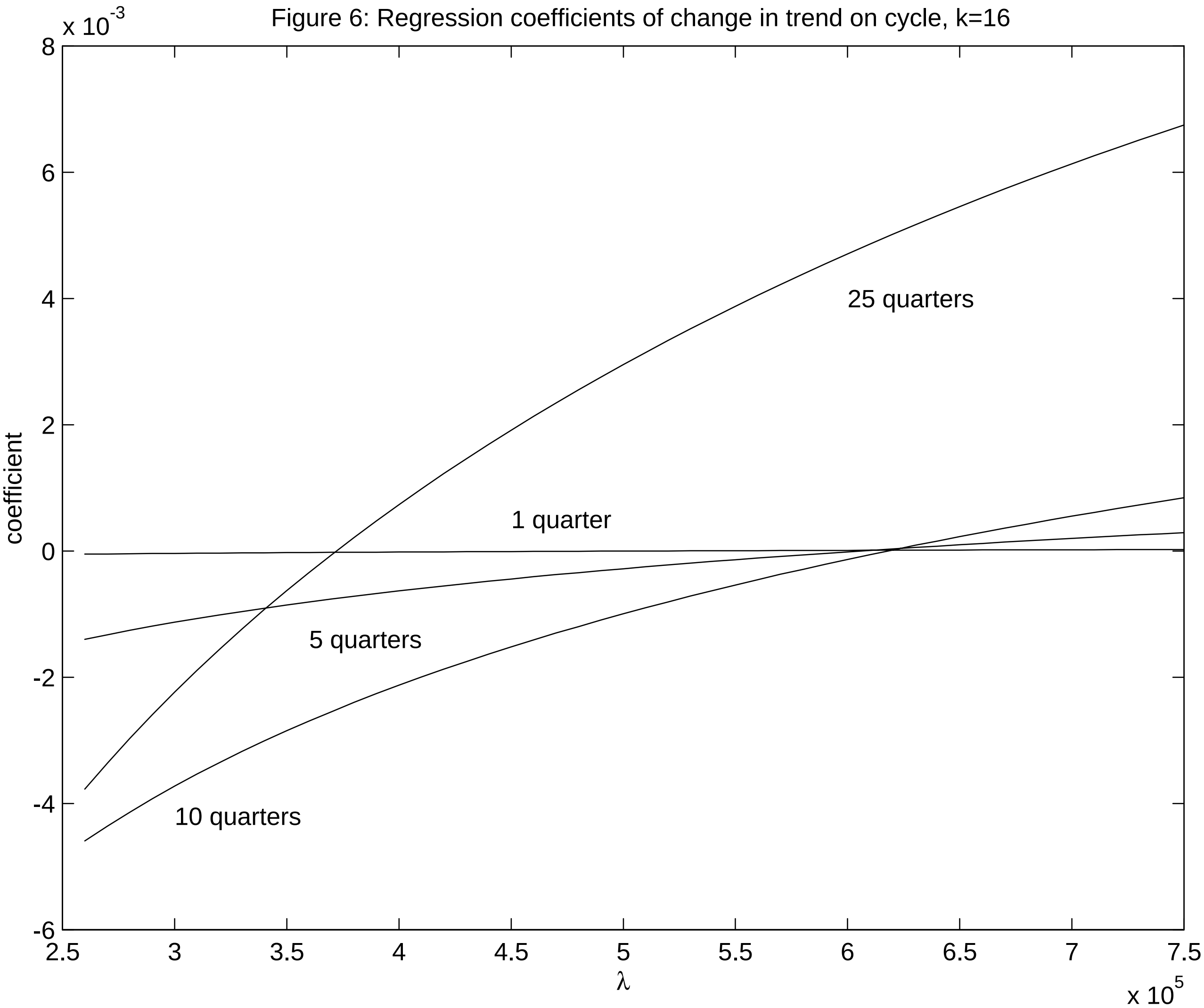

Figure 7: Regression coefficients of change in trend on cycle, y^{32}, k=12

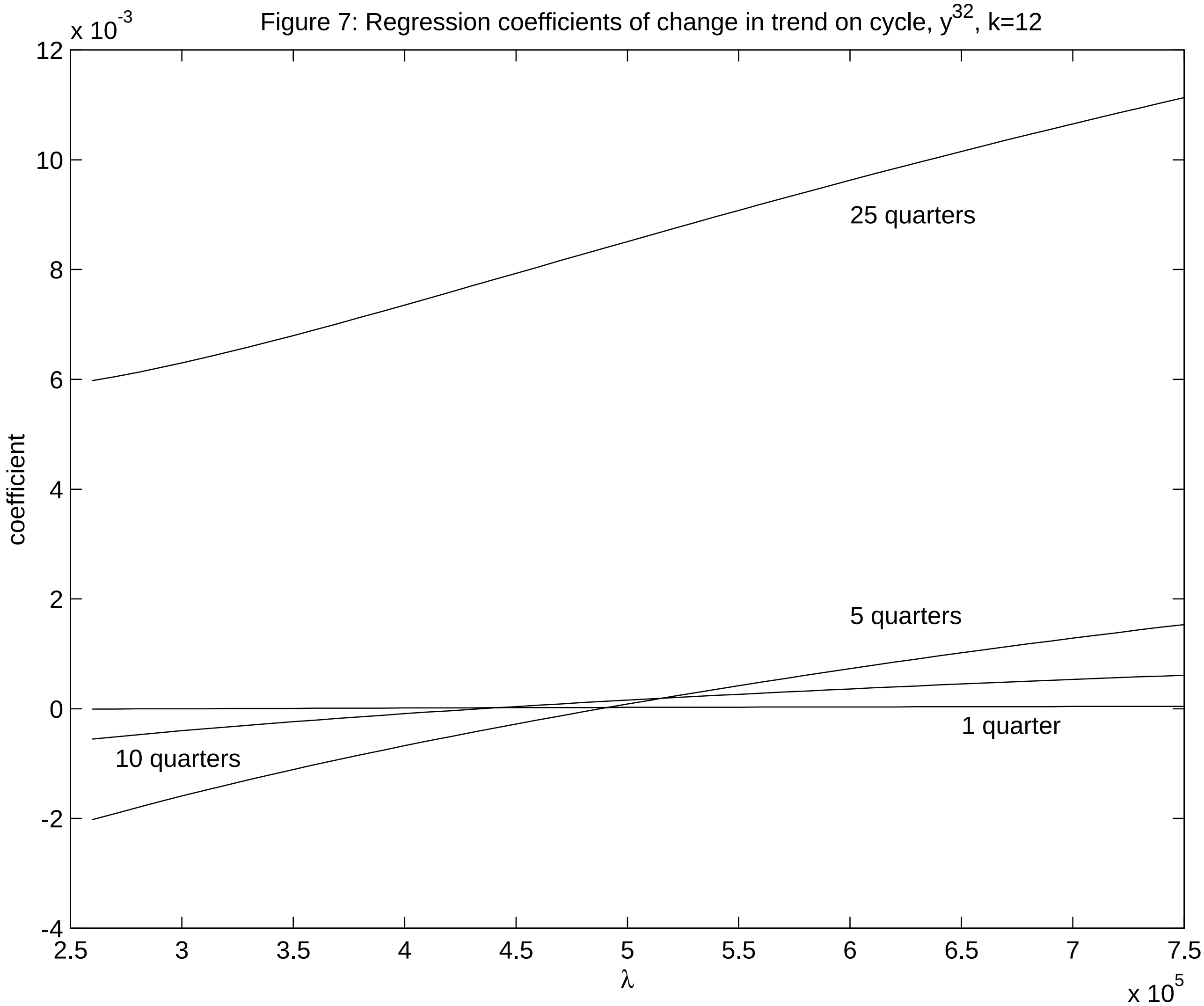

Figure 8: Gains for Various Filters (Values of λ in Parenthesis)

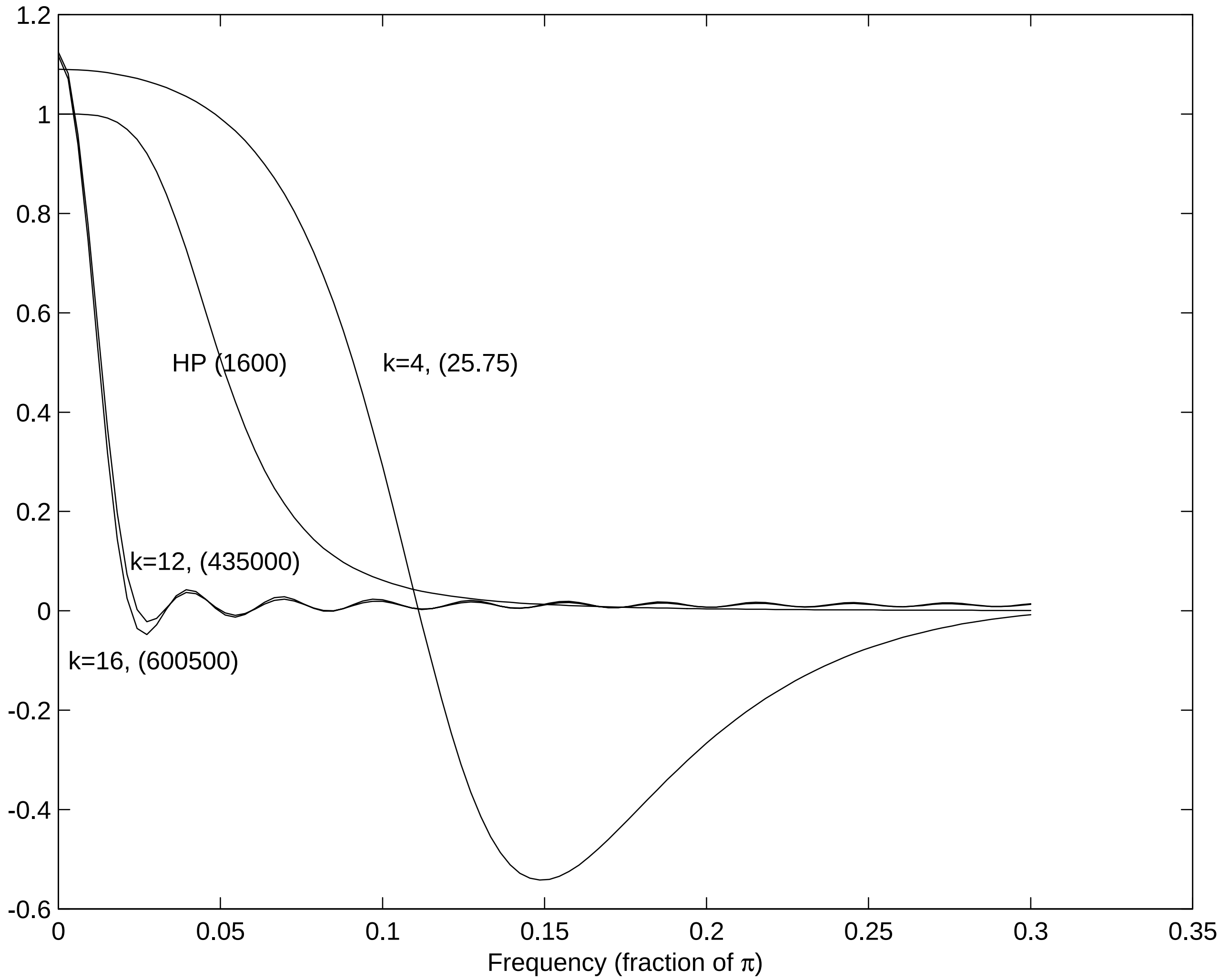

Figure 9: Regression coefficients of change in trend on cycle, y^{32}, k=4

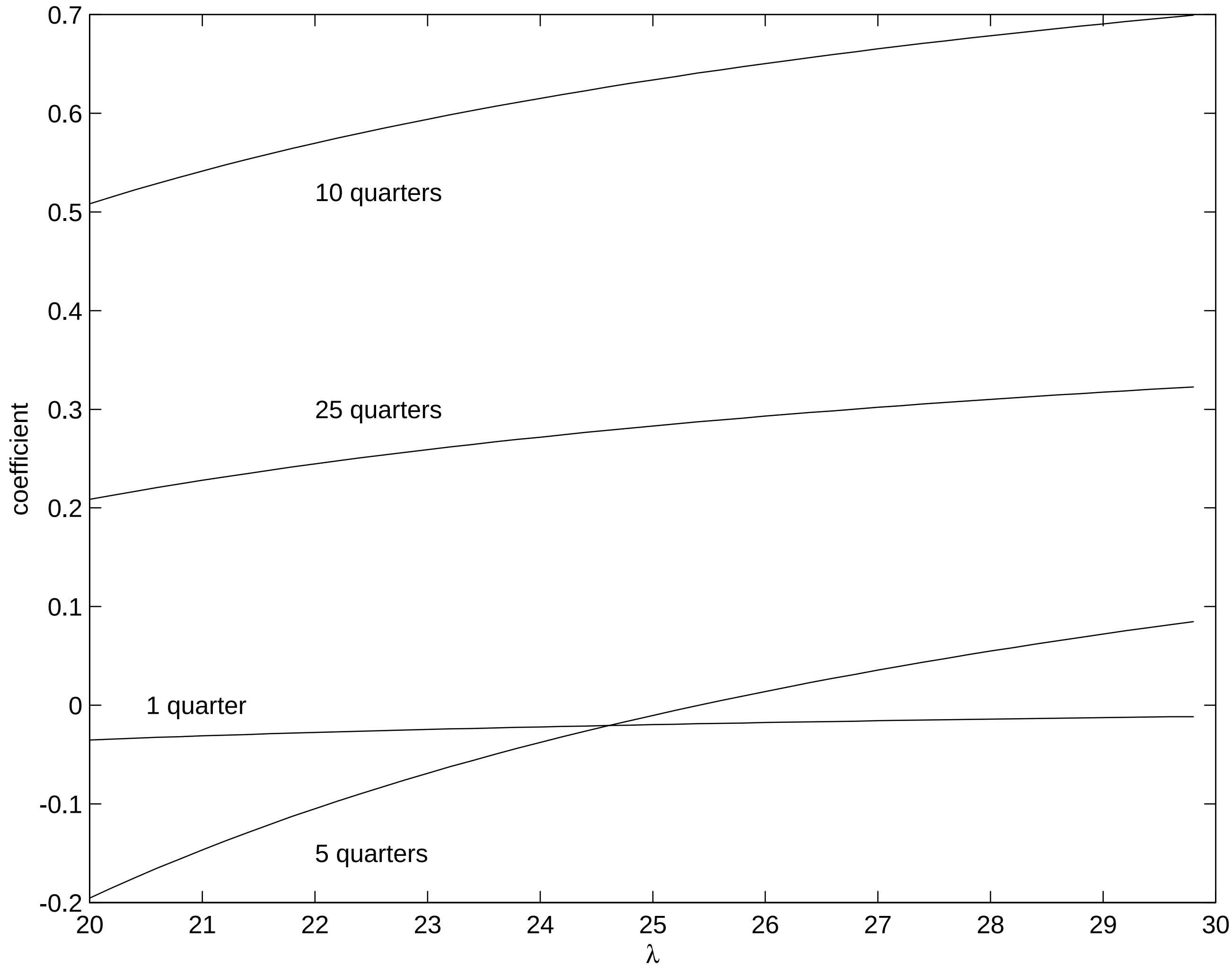

Figure 10: Path of y^{32} and its trend with k equal 4 and 16

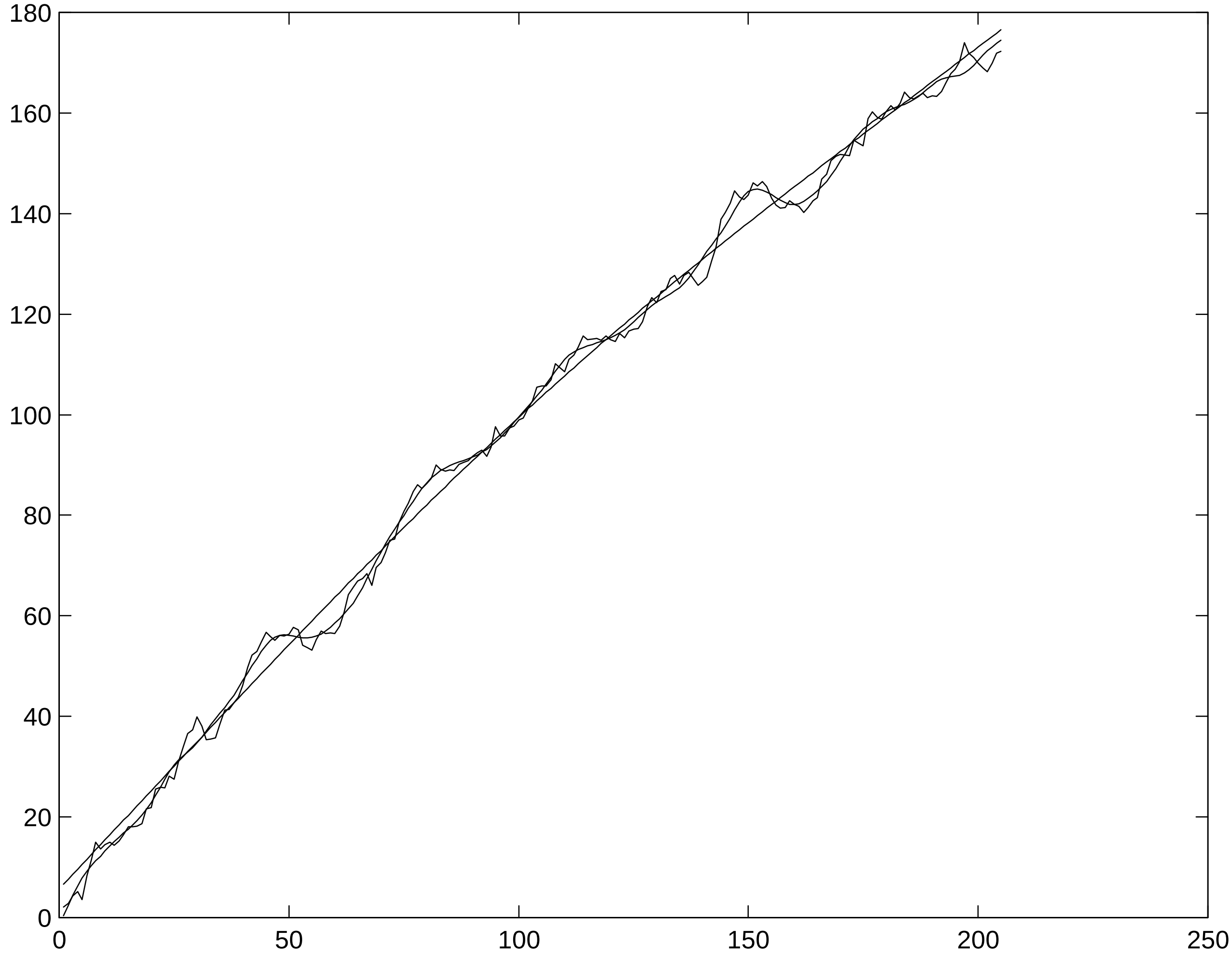

Figure 11: Errors in estimating D^2 with k=16 and linear trends

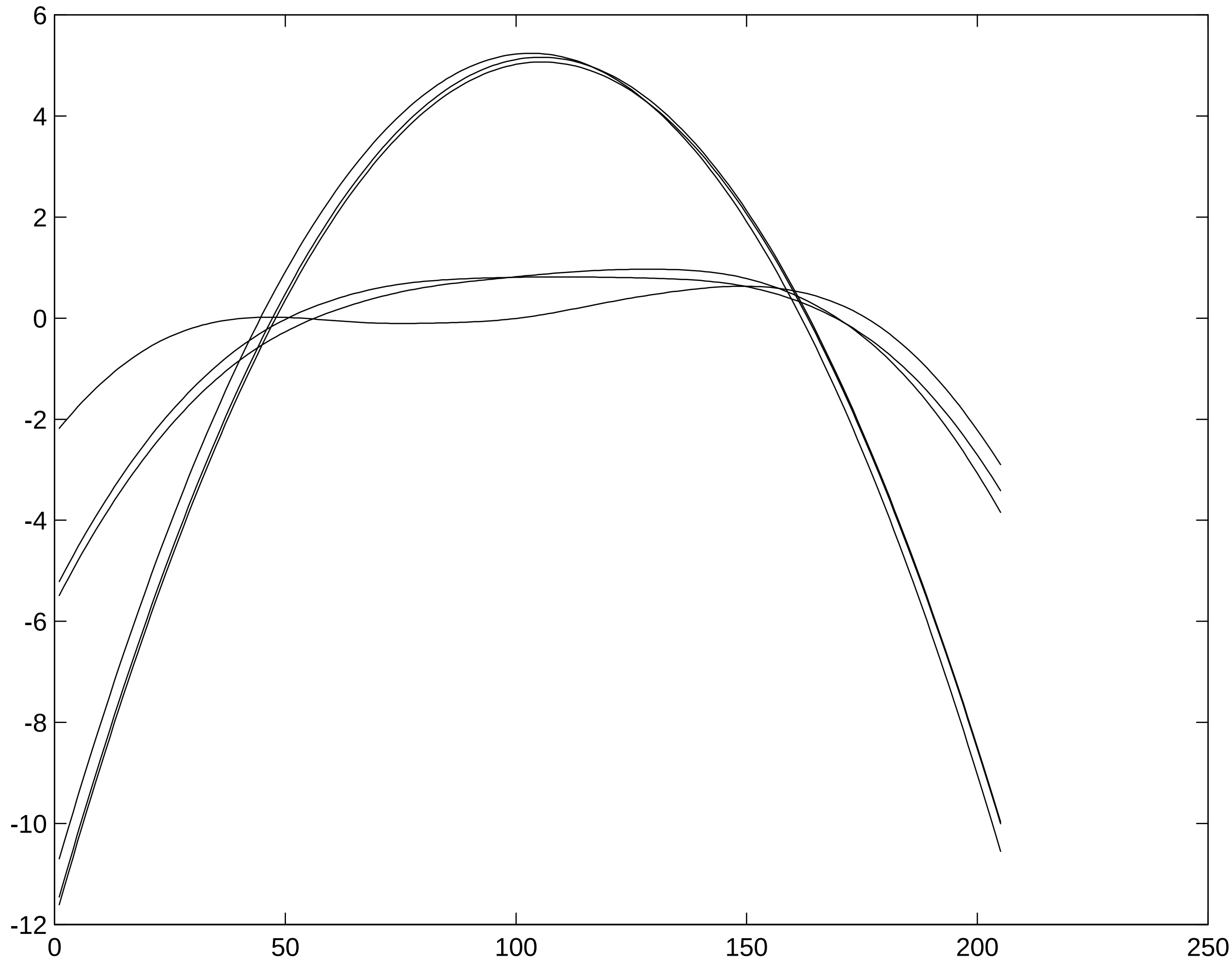

Figure 12: Errors in estimating D^4 and D^5 with k=16 and linear trends

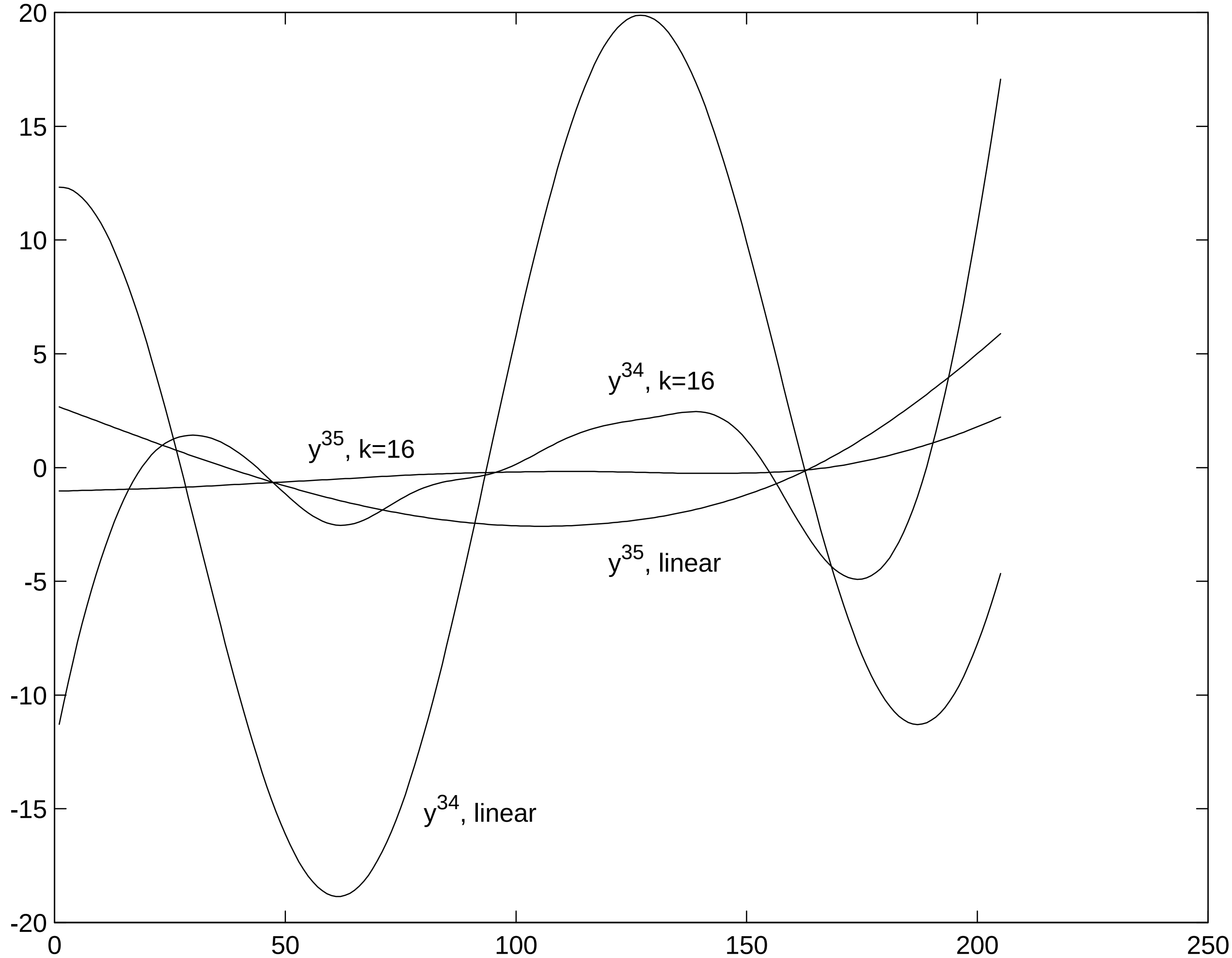

Figure 13: Errors in estimating C^1, C^2 and C^3 with HP filter

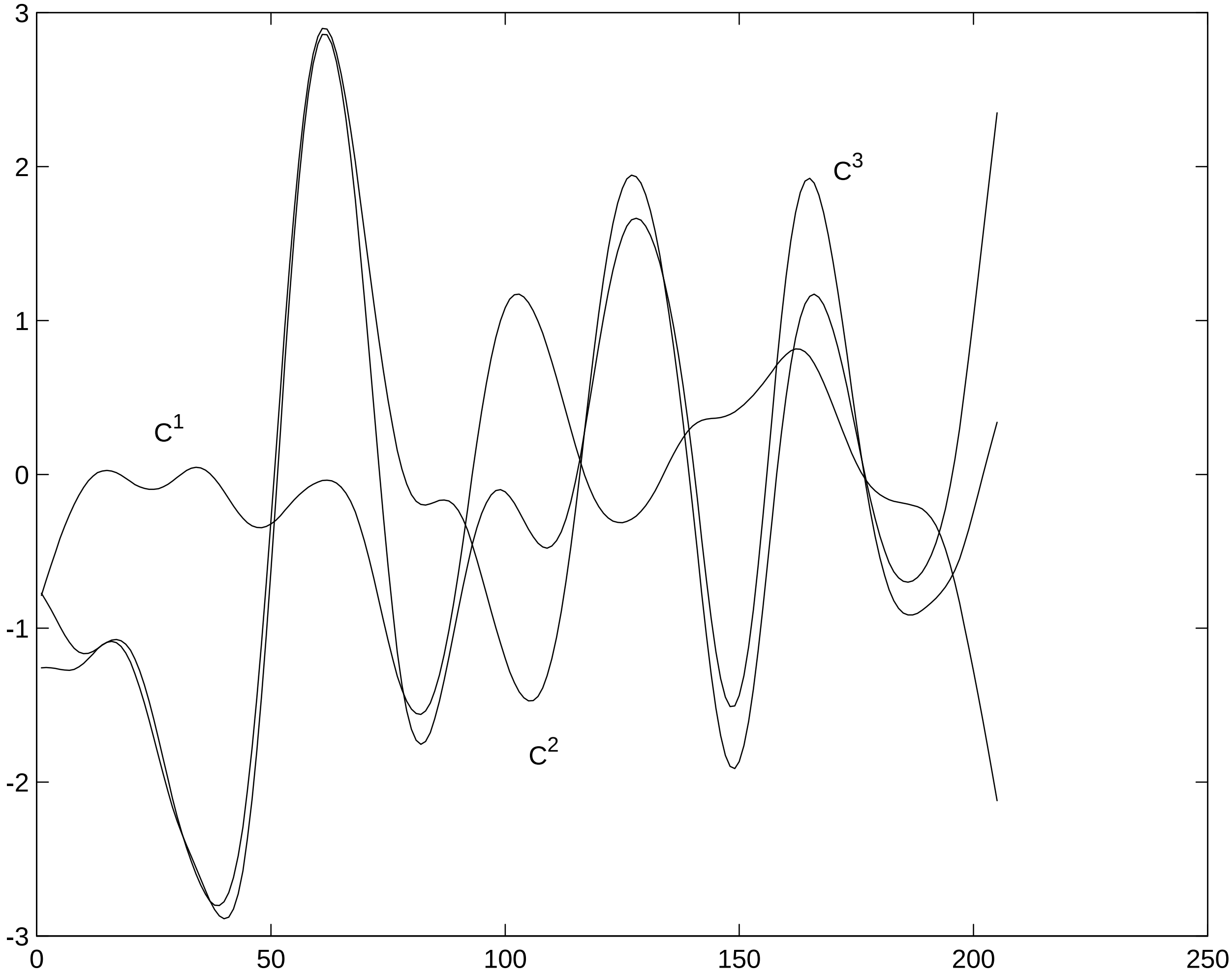

Figure 14: Response of cyclical y^{33} to shock of cyclical y^{12}

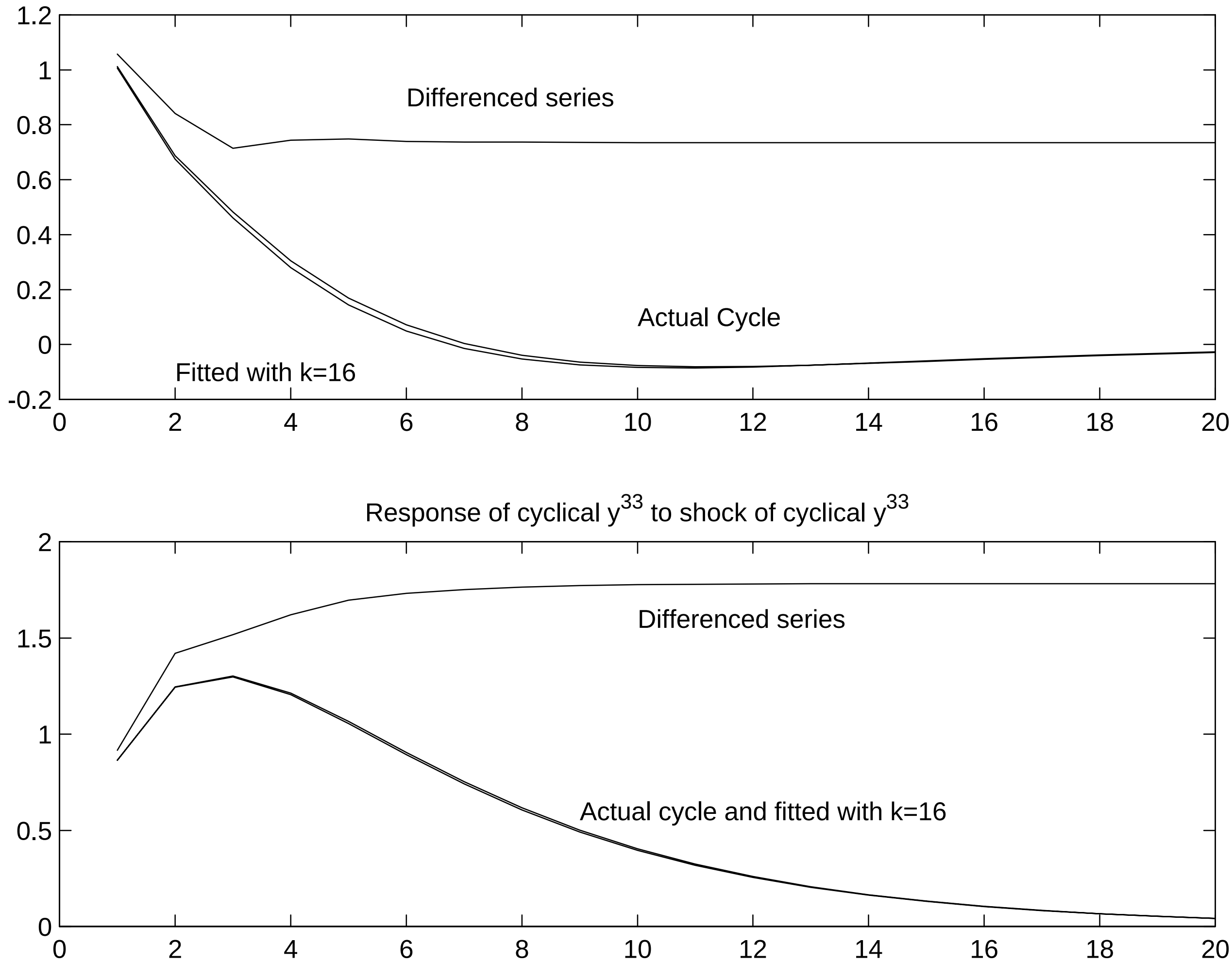

Figure 15: The Smooth Trend of U.S. GDP