ESTIMATION OF THE CYCLICAL COMPONENT OF ECONOMIC TIME SERIES

Maria-Helena A. Dias
Department of Economics of the State University of Maringá, researcher of the National Council of Research and Development (CNPq)
Joilson Dias
Department of Economics of the State University of Maringá, researcher of the National Council of Research and Development (CNPq)
Charles L. Evans
Department of Economics of the State University of Maringá, researcher of the National Council of Research and Development (CNPq)

Resumo

Este trabalho tem por objetivo a suavização de séries temporais com simulações de Monte Carlo para a análise de séries que apresentam mais de uma possível quebra estrutural, sejam estas advindas de movimentos no coeficiente da tendência ou no intercepto. O filtro de Hodrick-Precott (HP) não proporciona a identificação de tais mudanças na série, de intercepto e/ou de coeficiente, para computar e portanto limpar a série da tendência para que apenas o componente cíclico seja analisado. Quando as séries são relativamente estáveis, como é o caso de séries trimestrais de produtividade e emprego das economias desenvolvidas, esta característica do filtro HP não tem maiores implicações. Porém, para economias relativamente instáveis este ponto se torna relevante, pois a incidência de mudanças na tendência se torna maior, e o filtro HP pode levar os empiricistas a tratarem a suavização das séries de forma simplista. Dentro do exposto, propomos uma metodologia que considera a possibilidade de quebra em qualquer momento no tempo, seja de coeficiente ou de intercepto. Como exemplo, utilizamos modelos com variações na tendência, em coeficiente e em intercepto, dentro de uma metodologia recursiva de suavização de tendência para series obtidas através de simulações de Monte Carlo. Contudo, comparamos ainda uma aplicação para o PIB brasileiro.

Palavras-chave: ciclos econômicos, filtro hodrick-prescott, modelos de séries temporais.

Abstract

The objective of this paper is to show an alternative technique to smooth time series from Monte Carlo Simulations. The technique considers that time series can contain more than one structural break, coming from movements in coefficients of trend or from intercept. The Hodrick-Prescott Filter (HP) does not provide identification of such possible breaks in order to smooth trend from the series to analyze its cyclical component. If the series are relatively stable, this problem may not have relevant implications. Otherwise, for economies relatively unstable, trend movements may interfere in the specification of the cyclical component, and Hodrick-Prescott smoothing could lead empiricists to achieve simplistic forms to economic cycles. In the context, we present an empirical methodology that allows structural breaks in any point of time, from coefficients or from intercepts. We apply this recursive technique to different models with variations in trend, from coefficients and from intercepts, using series simulated by Monte Carlo. Moreover, we compare the results of both techniques to the Brazilian GDP.

Key words: business cycles, hodrick-prescott filter, time series models.

ESTIMATION OF THE CYCLICAL COMPONENT OF ECONOMIC TIME SERIES

Maria-Helena A. Dias*
Joilson Dias*
Charles L. Evans*

I. Introduction

The *objective* of this work is to present a technique to separate the cyclical component of a time series. A time series (Y) contains four basic elements, such as: the seasonality (S), the trend (T), the cyclical (CL), and the stochastic component (ER). Once one eliminates the seasonality and the trend of the variable time series, it remains with the cyclical and stochastic components. Hence, the main goal is to separate from the remaining part of the series, the cycle from the stochastic component.

The underlying *hypothesis* of the methodology proposed in this work is that the business cycle of an economic series may not necessarily contain the stochastic component error. This component could account for others sources of disturbances not essentially related to cycles or be completely random. One possibility is that this error could represent the first impact of shocks and would not demonstrate any cyclical behavior, such as a right away jump that stands for one period only. If one accepts that the random error of a series is not necessarily part of its cyclical component, one *question* rising is if the Hodrick-Prescott filter is able to capture the most approximated trend estimation as a tool to identify the true cycle of the variable, especially in cases of less stable time series.

Business cycles were a major *topic of discussion* especially in periods of large fluctuations in aggregate variables of industrialized economies. One known example is the Great Depression. Many researchers devoted themselves to explain the main cause of fluctuations and to understand their consequences to real economic activities. In 1946, an important work by Burns & Mitchell brought results about the measurements and classification of business cycles. Since then, the techniques to determine business cycles improved substantially. In the 1980s, with the diffusion of Hodrick-Prescott Filter (HP Filter) to smooth trends, new efforts took place toward the analysis of business cycles of modern market economies. The business cycle has since then been seen as deviations of trends of series, considering that these trends can change over time.

In accordance with all four components of a series, the introduction of old and new methodologies to smooth trends took place. There are some results using moving average techniques (Burns & Mitchell, 1946), some working with the hypothesis of changes in the growth rate of the variable for different periods of time (Hodrick & Prescott, 1980, 1997), some considering the logarithmic differences of the variables, and others constructing structural time series models to isolate the cyclical component (Harvey & Jaeger, 1993). Little attention came towards the stochastic component of the series, in the sense that it is a residual component with erratic behavior, which may bring no especial economic meaning. In fact, to Hodrick & Prescott (1980), Nelson & Plosser (1982), and theirs followers, cycles are deviations from trend of a series, together with its stochastic component.

In this context, the *proposition* here is to improve with a technique that separates the trend and the cycle of a variable, through an estimation of the cyclical component of the series, considering that there can be many changes in trends, coming from its slope or intercept. These trend characteristics should be accounted for when computing the cyclical component of any series.

II. Growth and Cycles

In the 1990s, many authors called attention to the fact that growth is related to cycles in time series, in the sense that one affects the other. Therefore, one should consider the growth rate of

* Professors at the Department of Economics of the State University of Maringá, Brazil. Researchers of the National Council of Research and Development (CNPq), Brazil.
** Vice-President Economist of Macroeconomic Research at the Research Department of the Federal Reserve Bank of Chicago, USA.

variables to isolate the cyclical component (Cooley & Prescott, 1995). In line with this thought, determining the nature of the long term trends (Hamilton, 1989) and detrending techniques became important to the analyses of business cycles (Harvey & Jaeger, 1993).

Hamilton (1989) develops a nonlinear iterative filter based on an algorithm that determines shifts in regimes, from a state of positive to negative growth rate. These changes describe business cycles fluctuations. The main hypothesis of the study is that a nonstationary series can have discrete shifts in its mean growth rate. As pointed out by Hamilton (1989, p. 358), the proposed algorithm is the statistical identification of "turning points" of a time series, which distinguishes from other rather arbitrarily techniques by specifying the "turning point" as *a structural event that is inherent in the data-generating process.* Thus, the shifts are represented in a model of trend with certain probability to switch regime, as in a Markov process.

In accordance with Hamilton's results, his Markov model can be used as an alternative method to NBER dating of business cycles.

According to Harvey & Jaeger (1993, p. 231), *mechanical detrending based on the Hodrick-Prescott filter can lead investigators to report spurious cyclical behavior*. This is especially true when H-P filter is applied to other economic series different from the US economy (1993, p. 236). For instance, Harvey & Jaeger found that the HP filter resulted in more volatile cyclical component to the real Austrian GDP, probably because erratic movements were important in the Austrian data. The same may happen to models based on autoregressive integrated moving average (ARIMA) with relatively small sample sizes. When sample is not long enough to allow the researcher to observe real facts of the series, one could specify a process being integrated of order 1 when in truth it was integrated of order 2. Moreover, the authors' advice is that a *proper presentation of the stylized facts associated with a trend plus cycle view needs to be done within the framework of a model that fits both components at the same time.* (Harvey & Jaeger, 1993, p. 246)

These developments of business cycles theories and methodologies were accomplished by evolutions in the measurements of trend changes, such as: Phillips (1987), Perron (1988), Phillips & Perron (1988), Perron & Vogelsang (1992), Vogelsang (1997), among others. The main issue of these works was to identify if the trend was stochastic or stationary. Then, they are useful to help the studies of detecting possible breaks in time series with changes in its intercept or slope. Sometimes a stationary time series has a unique break in the intercept of its trend, changing its level once and for all, in the analyzed period. This is important to business cycles because it could be happening through an exogenous shock that generates short run fluctuations, as well as long term movements, in economic activity.

In line with the authors cited here, sources of *motivation*, this research uses the central idea of the methods of detecting trend breaks to account for possible changes in trend when smoothing the variable. Thus, we can smooth the trend of the variable to remain with a better estimate of the cyclical and irregular components of the series. Nonetheless, achieving a good approximation of the real measurement of the cycle implies removing the irregular component of the series, throughout its estimation.

III. Methodology of Measuring Business Cycles

The objective of this section is to explain a new technique of accounting the cyclical component of the variable. The motivation of it came from analysis of HP filter of smoothing trend and from the methodologies of determining trends and detecting their possible breaks.

Considering the four elements of a time series, once adjusted for seasonal fluctuations, the series remains with the trend (T), the cycle (CL), and its irregular (ER) component. The trend of the variable can be deterministic or stochastic. If the trend has a deterministic behavior, it would be relatively easy to estimate it and to smooth the series from its possible effects to get an approximation of the real cyclical component. However, if the trend is stochastic, it is a relatively harder matter to determine its form. In this work, we are considering that a stochastic trend could have more than one shape, such as: additive, multiplicative, with constant, without constant, to cite a few. Thus, after identifying the trend of the series, the cyclical component and its irregular variability would remain yet. For the series that do not vary much, probably the stochastic term would not affect the cyclical component strongly. Although we would still have to account for it in order to find an appropriate estimation of the cyclical component. Achieving this appropriate estimation of the cycle imply the application of a technique to smooth the true trend and to estimate at least the path of the irregular component. Our method for doing so will be displayed thereafter, following a Monte Carlo procedure.

3.1. Monte Carlo Experiments

The simulations we are about to present here follow Maddala & Kim (1998)'s explorations on deterministic trend and stochastic trend. Their work emphasizes detrending. Thus, the cyclical component they specify is equal to $-0.3ER_t$. Moreover, it is supposed to be a *mean-zero stationary process*. The authors chose to apply a model with no autocorrelations in the error term, ARIMA (0,1,1).[1]

In what follow, we explore the experiments used to explain the technique to isolate the cyclical component of a series, considering the possible occurrence of stochastic trends and the erratic term of the variable.[2]

3.1.1. Model I

Initially we created four models of time series with different features for each of their components. In the first model, Model I, we made a time series (Y_1) with one hundred observations (n=100), in which its components have the following characteristics:

$$\begin{aligned} &Y_{1t} = DT_{1t} + ST_{1t} + CL_{1t} + ER_t, \\ &DT_{1t} = 1+0.5*TREND, \\ &ST_{1t} = TREND*ER_t, \\ &CL_{1t} = -0.5*ER_t, \\ &ER \sim N(0, 0.5^2). \end{aligned} \tag{1}$$

Where DT is the deterministic part of the trend, ST represents the stochastic behavior of the trend, CL represents the cycle, and ER is the irregular component, it is expected to have a normal distribution with zero mean and 0.25 value for its variance. As usual, the trend is a variable that assumes the value of 1 for the first observation, 2 for the second, and so on to get equal to 100 to the last observation. The subscript t is to indicate the time. The series generated by this Monte Carlo procedure did not contain the seasonal component.

Moreover, estimating possible changes in the trend of Y_1 implies a regression analysis in which the regressors can capture such behavior. Our regression model includes dummy variables for this purpose. Therefore, we constructed a hundred dummy variables to account for possible changes in the intercept of the Y_1. We named it as DC1, DC2, …, DC100. DC1 is represented by a vector of 1 by 100 that has the following form: [1, 1, 1, …, 1] . DC2 is a vector of 1 by 100 with the form of [0, 1, 1, …, 1]. DC3 has the form of [0, 0, 1, …, 1]. DC99 is the following: [0, 0, 0, …, 1, 1]. And finally, the dummy variable DC100 is [0, 0, 0, …, 1]. Using the same logic, we also built a hundred dummies to represent possible changes in the slope of the trend of the time series, Y_1. These dummies received the names of D1, D2, D3, …, D100. The first dummy variable D1 has the following form: [1, 2, 3, 4, 5, …, 100]. D2 is [0, 1, 2, 3, 4, …, 99]. D3 follows the vector [0, 0, 1, 2, 3, …, 98]. After all dummies, we have D100 as [0, 0, 0, 0, 0, …, 1]. In addition, we considered a variable lagged on time by the order of one, such as: Y_{1t-i}, where i=1. It is worth to mention that we run a lag test to determine the size of i.

Hence, the regressions we estimated to account for possible changes in the trend in order to smooth our Y_1 series, in a recursive manner, according to the following rationality:

[1] Maddala & Kim (1998, pp. 29-34).

[2] The mathematical representations of the simulations we developed are displayed in the appendix.

$$
\begin{aligned}
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + ER_t \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D2_t + \beta_4 DC2_t + ER_t \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D3_t + \beta_4 DC3_t + ER_t \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D4_t + \beta_4 DC4_t + ER_t \qquad (2) \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D5_t + \beta_4 DC5_t + ER_t \\
&\vdots \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D99_t + \beta_4 DC99_t + ER_t \\
Y_{1t} &= \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{1t-1} + \beta_3 D100_t + \beta_4 DC100_t + ER_t
\end{aligned}
$$

In accordance with the set of regressions above, we estimated at least one regression for each observation of our variable Y_1, using the Ordinary Least Square technique, in a recursive manner. Notice that the first regression is the simple model with a constant and a linear deterministic trend. Maddala & Kim (1998, p. 70) stressed this model to analyze its asymptotic distribution.

In order to hold consistent estimates of the β parameters, we considered only the statistical significant results at the level of 90 percent of confidence. For the cases that β_3 and β_4 were not significant, we considered the estimates of the first equation and thus, we reported them for that case (observation). With the results of these calculations, we were able to construct a residual series for each regression, considering that there could be changes in trends coming from movements in the constant term or in its slope in each point of time during the period in analysis.

The results were then gathered into a (99x99)[3] table with a series of residuals of each significant regression, one at each column, considering possible changes in trends coming from intercept or slope movements. Thus, each observation had an estimated residual that goes through each of the regressions. For example, this implies that the observation 20, t=20 at the horizontal line of the table, had 99 estimates of the regressions residuals. Once the table of residual series was ready, we chose to delete the first four and last four observations of residuals because we understood that they could have statistical problems in their estimates. This is possible because the dummy variables turned out to be mostly one for the initial four and zeros for the last four observations. Since the dummies for the observations under numbers 2, 3, and 4 reached a near singular matrix, we decide to repeat the results for the regression of the first observation in these series.

In addition, we computed the mean of residuals for each observation, to get a vector of the mean of residuals, calculated by each regression, containing the cyclical and the irregular components of the variable, with 92 observations. Moreover, we also computed the standard deviation of the residuals of the regressions for each individual observation, coming up with a path for the residuals, containing the cycle and the error of the variable Y_1.

This procedure allows us to construct a path along with the cyclical component of the series could vary into, considering the standard deviations computed through the residuals of the regressions applied for Y_1.

The main objective of this experiment is identifying the cyclical component of the series, despite the behavior of its trend and irregular components. Getting the closest approximation of the true cycle of the series implies other steps that we will show later on. Since ER is erratic or stochastic, we expect to be able to find only a reasonable estimation of its part through the residuals obtained in the regressions of Y_1.

The Monte Carlo process implies we have in advance the composition of the cyclical component of this series. Thus, the idea is to obtain an estimate as close as possible of the true cyclical component.

[3] Recall variable Y_{1t-1}, this was significant for most part of models we estimated.

Initially we thought to use the mean of the residuals computed to each observation (MRES) minus its standard deviations (STDRES). However, *these are second order deviations and so they are all positive*. Therefore, this simple procedure may not be able to capture the most approximated estimation of the cyclical component of the series. It would probably have a bias in the results. Nonetheless, we tried this procedure and the result was the opposite of the true cycle, indicating that a better estimation would be to *sum* the standard deviation of the residuals (STDRES) to their mean (MRES), and then to multiply the result to (-1). Fortunately, our estimated cyclical component showed to be close to its true value, generated by the Monte Carlo process. The equation to represent the technique to determine the cycle is as follows.

$$CL_{1t} = (MRES_{1t} + STDRES_{1t}) * (-1) \qquad (3)$$

Figure 1 displays this result. As one may notice, the estimated cycle mimics the cycle of this Monte Carlo model for most part of the sample. There are throughout Figure 1 some deviations of the true cycle for the years of 78 to 81.

The main question remained was if the procedure proposed here would hold to other models whose time series have different trend functions. This investigation will take place through the simulation of three extra variables Y_2, Y_3, and Y_4, our models II to IV. All models where also estimated using the popular technique of Hodrick-Prescott filter in order to give us a parameter to compare with. For that a specific section will show the evaluation of the main results between the two techniques.

3.1.2. Model II

The second model, Model II, we generated with a different pattern to the stochastic trend of the variable. The main objective was to test the procedure to find the true cyclical component of the variable.

The time series (Y_2) has n equal to 100 observations, and its components are the deterministic trend (DT), the stochastic trend (ST), the cycle (CL), and the irregular component (ER). These components have the following characteristics:

$$\begin{aligned} &Y_{2t} = DT_{2t} + ST_{2t} + CL_{2t} + ER_t, \\ &DT_{2t} = 1+0.5*TREND, \\ &ST_{2t} = [1+(0.5*TREND*ER_t)] / [1+(0.5*TREND)], \\ &CL_{2t} = +0.25*ER_t, \\ &ER \sim N(0, 0.5^2). \end{aligned} \qquad (4)$$

ER has normal distribution with mean zero and variance 0.25.

Notice that the stochastic trend is more complex than that of the first model. The purpose here is finding out if the cyclical component of a relatively more erratic variable could be specified by the new technique that we are proposing.

In addition, applying the new technique implies in exercising with the same dummy variables we used before because the number of observations is the same. The regression to estimate the parameter values of the components of Y_2 also followed those to calculate Y_1 cyclical component. Therefore, the basic regression repeats those of the first model with the applications to Y_2, that is:

$$\begin{aligned} &Y_{2t} = \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{2t-1} + ER_t \\ &Y_{2t} = \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{2t-1} + \beta_3 D2_t + \beta_4 DC2_t + ER_t \\ &\quad \vdots \\ &Y_{2t} = \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{2t-1} + \beta_3 D99_t + \beta_4 DC99_t + ER_t \\ &Y_{2t} = \beta_0 DC1_t + \beta_1 D1_t + \beta_2 Y_{2t-1} + \beta_3 D100_t + \beta_4 DC100_t + ER_t \end{aligned} \qquad (5)$$

In accordance with our new technique, all the procedures applied in Model I was used in Model II to reach the estimates of the parameters β_j, j=0,1,2,3,4.

One different result in Model II was the estimate of β_2. As one should expect, none of the regressions had significant estimate for this parameter. Despite of this, the main result about the estimation of the cyclical component holds. We can isolate the cyclical component by calculating the mean of the residuals of each regression applied to each observation at the overall period. Then, estimating the standard deviations of these means gives a path of residuals containing both cycle and irregular terms. In addition, the most approximated estimation of the true cyclical component of the variable Y_2 resulted from the following formula:

$$CL_{2t} = (MRES_{2t} + STDRES_{2t}) * (-1) \quad (6)$$

Where MRES represents the mean of the residual of the regressions of each observation containing the dummies, accounting for changes in trends, and STDRES is its estimated standard deviation. Figure 2 shows the resulting cycles to this model. It is worth mentioning that the estimations of this model are not as accurate as those of model I. Roughly, the smoothing technique seems still giving a good approximation of the true cycle. We will come back to this matter later on.

Hence, the smoothing technique using dummies to account for possible changes in trend in each period combined with the technique to extract the error term of the variable to estimate the cyclical component constitute a new methodology of analyzing business cycle.

3.1.3. Model III

In this model, we constructed a variable with a unique and permanent break in the stochastic trend. The objective was to test our methodology robustness in estimating the cyclical component of the variable, isolating its trend throughout with its smoothing. The procedure to achieve the estimated cycle was the same as before but the variable (Y_3) differs substantially.

$$\begin{aligned} &Y_{3t} = DT_{3t} + ST_{3t} + CL_{3t} + ER_t, \\ &DT_{3t} = 1+0.5*TREND, \\ &ST_{3t} = 1+(0.5*TREND), \text{ for } t=1,2,\ldots,50, \\ &ST_{3t} = 0.5*TREND, \text{ for } t=51,\ldots,100, \\ &CL_{3t} = -0.25*ER_t, \\ &ER \sim N(0, 0.5^2). \end{aligned} \quad (7)$$

In this specification, we have that the trend of the variable Y_3, $T_{3t,}$ turns out to be the following:

$$\begin{aligned} &T_{3t} = 2+2*(0.5*TREND), \text{ for } t=1,2,\ldots,50, \\ &T_{3t} = 1+2*(0.5*TREND), \text{ for } t=51,\ldots,100. \end{aligned} \quad (8)$$

In line with this framework, we applied our calculations of the regression analysis to generate the residuals of each observation and to form one series with the estimated mean and another one of its standard deviations. In this model, the resulting cyclical element was also compatible with the formula used for models I and II. This is displayed in the sequence.

$$CL_{3t} = (MRES_{3t} + STDRES_{3t})*(-1) \quad (9)$$

The graphical representation of the estimated cycle for this model is in Figure 3. Yet, finding out the right specification to the cyclical component of the series implies that new simulations must be conducted. Thus, Model IV brings new insights into the research.

3.1.4. Model IV

The variable generated in this model, called Y_4, consists in a different feature than that of Y_3. This is because the cyclical component of Y_4 is not multiplied by a negative constant. Instead, Y4 has the following characteristics:

$$
\begin{aligned}
&Y_{4t} = DT_{4t} + ST_{4t} + CL_{4t} + ER_t,\\
&DT_{4t} = 1+0.5*TREND,\\
&ST_{4t} = 1+(0.5*TREND), \text{ for } t=1,2,\ldots,50,\\
&ST_{4t} = 0.5*TREND, \text{ for } t=51,\ldots,100,\\
&CL_{4t} = 0.25*ER_t,\\
&ER \sim N\ (0,\ 0.5^2).
\end{aligned} \tag{10}
$$

In this estimation of the cyclical component of series Y_4, the results indicate that the right formula to reach the true value of the cycle is different from the one used for models I, II, and III. Partially, the results of this fourth model do not confirm those findings of models I to III. In contrast, the best solution we got follows the formula bellow, using the mean and the standard deviations of the residuals of the regressions applied.

$$CL_{4t} = (MRES_{4t} + STDRES_{4t}) \tag{11}$$

Plots of the resulting estimated cycle of Model IV is in Figure 4.

The matter now is to discover when the solution to determine the cyclical component should use CL_{2t} or CL_{4t} formulas. Finding out the answer to this problem exhorts the exploration of the elements of this new methodology, such as: the behavior of the models' variances.

Beginning with this analysis, we compared two elements with the mean cycle (same as MRES), namely: the variance of the residuals of the regressions, and their standard deviations. The objective here is to find a condition or hint in the estimated parameters to guide us in choosing the equation that could calculate the true cycle of the variable. Thus, we tried to find some pattern in the parameters that could signalize the direction we should take to select an equation to calculate the cycle. The results are as follows.

For all four models, the respective estimated error, featured by the variance component of the residuals of the smoothing regressions (the square of the standard deviations), match exactly the mean cycle. Then, the results suggest that the variance of the residuals of the estimated regressions may have a distribution with zero mean and very small deviations. Figures 5 to 8 graph this condition. Although this is a characteristic corroborating to the statistical features of the methodology, it does not give us any direction on what formula to specify the cycle one should rely on.

In contrast, when we compare the mean cycle with the standard deviation multiplied by two (STDEV*2) for the four models, we have that their mean cycle are always higher than their STDEV*2. Figures 9 to 12 shows these results. Again, we find no hint on the true cycle formula. Because the models did not show any difference among them that could justify a different formula of calculating the cycle.

IV. Comparing HP Filter and DDME Methodologies

This section contrasts the HP filter for smoothing series as a tool to find the cyclical component of the series with the technique we are proposing here.

Considering the variable has two components, trend and cycle. Applying the HP filter to smooth the trend of Y_1 implies a smoothed series called y_1. This would give the value of Y computing with a smoothed trend. The difference between Y_1 and y_1 would give the residual component containing the cycle. Comparing the resulting HP cycle with the cycle generated by our Monte Carlo procedure requires creating a new variable with the remaining elements of the series after smoothing for its trend, we shall call this remaining part of the series of CTE_t, that is CL_t plus ER_t.

4.1. Model I

In accordance with the first model, the component CTE_t is CTE1. This corresponds to our variable MRES, or the mean of the residuals of the regressions with smoothed trend. We shall refer to it as the mean cycle of the variable.

According to Figure 17, one may observe that both estimations of the cycle for Model I (CTE1), HP-cycle1 and the mean-cycle1, are close to its value. However, testing their robustness in performing CTE1 requires other investigations through tests of mean, of median, and of variance.

Considering the t-test and the ANOVA F-statistics, the *tests of the mean* show that both statistics accept the null hypothesis that both series (mean-cycle1 and CTE1) has the same mean, the tests respective p-values are 81%. Then, the mean-cycle1 and CTE1 has the same mean. According to the mean test, HP-cycle1 also has the same mean as CTE1, with p-value of approximately 92%.

The *tests for equality of medians* between series show that the probability of mean-cycle1 belonging to the same distribution as CTE1 goes from 100% to 77%. Through the Wilcoxon/Mann-Whitney method, the test has the null hypothesis that the two series are independent samples from same general distribution. In this case, its probability is about 83%. The median Chi-square test is based on the rank of the observations of each series with an ANOVA test based on the comparison of the numbers of observations above and below the median of each series. In this test, the mean-cycle1 proves to come from the same distribution as CTE1 by the probability of 100%. The HP-cycle1 also passed in the test of medians. Its probability of belonging to the same distribution as the CTE1 series goes from 56% to 94%.

In accordance with the *tests for equality of variances* between series, the results indicate that the probability of equality ranges from 72 to 87% between mean-cycle1 and CTE1. The F-test here take the series with larger variance (L) and divides it by the series with the lowest variance, such that $F=(s_L^2/\ s_S^2)=(0.01917/0.01853)=1.0345$. The Bartlett test *compares the logarithm of the weighted average variance with the weighted sum of the logarithms of the variances* (EViews, p. 163). For this test, the adjusted statistic reported for series mean-cycle1 is 0.026 with probability of 87%. The Levene test performs *an analysis of variance (ANOVA) of the absolute difference from the mean* (EViews, p.163). This test indicates that the probability of the variances of the series mean-cycle 1 and CTE1 be equal is 78%.

For the case of the series HP-cycle1, the residual cycle of the HP-Filter, the probability for equal variance between this series and CTE1 ranges from 25 to 47%, much lower than the series mean-cycle1. Both F-test and Bartlett test for HP-cycle1 have probability of approximately 47%.

In addition to these results, we can show that the estimations for the cycle using the methodology DDME, to account for mean-cycle1, are better than the estimations with HP-Filter, for HP-cycle1, in Model I as follows. Since

$$\begin{aligned} &VAR(CTE1) = (0.138452)^2, \\ &VAR(DDME1) = (0.136123)^2, \text{ and} \\ &VAR(HP1) = (0.128497)^2. \end{aligned} \tag{12}$$

Then, we can compute how close the measurement of the cycle (mean-cycle1 or HP-cycle1) is to the cycle ($CTE1=CL_{1t}+ER_t$), resulting from Monte Carlo experiment. Therefore,

$$\left(\frac{0.136123}{0.138452}\right)^2 > \left(\frac{0.128497}{0.138452}\right)^2 \Rightarrow 97\% > 86\%. \tag{13}$$

Thus, the mean-cycle1 resulting from DDME method mimic the true cycle closer than the HP-cycle1, showing to be 11% better.

4.2. Model II

In spite of the results reported here, the estimations of Model II showed to be similar to those of Model I. Comparing the components of Figure 18, it suggests that mean-cycle2 is closer to CTE2 that HP-cycle2. The *tests for equality of mean* imply that the probability of mean-cycle2 and CTE2 have same mean is 92%. Moreover, this probability for HP-cycle2 series is approximately the same.

In Model II, the probability of mean-cycle2 has the *same general distribution* as CTE2, yet different samples, goes from 88% to 100%. The highest resulting p-value comes from Median Chi-square statistics. Again, the HP-cycle2 passes the test of same median as CTE2, showing p-values from 77 to 92%. However, these p-values are still lower than those found to the series mean-cycle2.

The *tests for equality of variance* show lower probabilities than those of Model I. Although the results for the series mean-cycle2 are significantly higher than the ones calculated for series HP-cycle2. In Model II, the probability of mean-cycle2 has the same variance as CTE2 ranges from 36 to 49%. If one considers the series HP-cycle2, its probability to have same variance as CTE2 ranges from 8% to 11%. This last result suggests the rejection of the null hypothesis of same variance of HP-cycle2 as CTE2.

Despite of this result, considering the variances of CTE2, of mean-cycle2 calculated through the residuals of the DDME smoothing procedure, and HP-cycle2 resulted from the application of the HP-Filter on the Y_{2t}.

$$\begin{aligned} &VAR(CTE2) = (0.207678)^2, \\ &VAR(DDME2) = (0.188464)^2, \text{ and} \\ &VAR(HP2) = (0.176366)^2. \end{aligned} \tag{14}$$

Therefore, we show the following comparison of variances:

$$\left(\frac{0.188464}{0.207678}\right)^2 > \left(\frac{0.176366}{0.207678}\right)^2 \Rightarrow 82\% > 72\%. \tag{15}$$

The result of the comparison of the variances indicates that the mean-cycle2 is 10% closer to the movements of CTE2 than is HP-cycle2.

4.3. Model III

Turning the attention to the third model, not all the results hold. Getting inspection on Figure 19 suggests that the resulting HP-cycle3 is closer to the true cycle CTE3 than is mean-cycle3, especially to the period after the break in the series Y_{3t}, exactly in the 50^{th} observation. Nonetheless, we proceed with the study of the methods HP-Filter and DDME smoothing.

Comparing the *tests for equality of means*, the p-value for equality between series CTE3 and mean-cycle3's means is approximately 97%. For the case of HP-cycle3 we have the probability of 93% of being equal to CTE3, which is highly acceptable.

Moreover, the *tests for equality of medians* of two series indicate that the probability of mean-cycle3 has the same distribution as CTE3 ranges from 77 to 88%, but independent samples. The best p-value is given by Adjusted Median Chi-square statistic, which is a continuity corrected statistic. The p-values for the series HP-cycle3 goes from 77 to 91%, it achieves a highest probability of belonging to the same general distribution as CTE3.

In addition, the *tests of variances* indicate that the probability of mean-cycle3 has the same variance as CTE3 can reach 70%, but it is 1.6% to the F-test. This may be happening because of the break in series Y_{3t}. Yet it has puzzled us since we expect the DDME method be a good smoother when the series has a break. Maybe we could have better estimates if we improve the econometric methodology to regress the variable against its trend from OLS to a more sophisticated one. The HP-cycle3 reaches higher probability of having same variance as CTE3, 90%, however it also gives a low p-value to the F-test, 27%.

In this context, the comparison of amplitude of the cycles computed by DDME and HP methodologies are the following:

$$\begin{aligned} &VAR(CTE3) = (0.207678)^2, \\ &VAR(DDME3) = (0.268027)^2, \text{ and} \end{aligned} \tag{16}$$

$VAR(HP3) = (0.232822)^2$.

Then, these variances imply that:

$$\left(\frac{0.268027}{0.207678}\right)^2 > \left(\frac{0.232822}{0.207678}\right)^2 \Rightarrow 166\% > 126\%. \quad (17)$$

This means that DDME is overshooting the variance of the cycle by 66%, then the estimated series mean-cycle3 has a relatively higher amplitude than CTE3. The HP methodology also overestimates the real cycle, but in a smaller amount, 26%.

4.4. Model IV

The last model to analyze is Model IV. Recall that this model also has a break in its trend as Model III but its cycle component is equal to $0.25ER_t$, and not equal to its negative, $-0.25ER_t$, as in Model III.

According to Figure 20, comparing results for plotted cycles, it is possible to notice that mean-cycle4 and HP-cycle4 have similar performance to mimic CTE4 until it reaches observation 50, where we simulated the break in the series. After that, it seems HP-cycle4 gets closer to CTE4. Let us investigate the statistical results.

The estimates for the *equality of means* are close to both methods but DDME is still better off in this test. Then, the probability of mean-cycle4 to have the same mean as the CTE4 series is approximately 94%, while this p-value for HP-cyle4 is about 92%.

When analyzing the results for the *test of equality of medians* between series, the estimates show that the probability of mean-cycle4 belongs to the same general distribution as CTE4 ranges from 85 to 100%. In this test, the HP-cycle4 goes from 77 to 96% of having same distribution as CTE4.

In the *analysis of the variances* and their possible equalities, the results show that the probability of mean-cycle4 series have the same variance as CTE4 ranges from 25 to 95%. HP-cycle4's p-values also go from 27 to 98%, but in different test-statistics. The F-test of mean-cycle4 has p-value of 25% and of HP-cycle4 of 98%. The Siegel-Tukey test indicates a p-value of 95% for mean-cycle4's variance and of 27% for HP-cycle4's variance. This Siegel-Tukey test assumes that the series have equal median but are independent from each other, which is the case in both pair of series, mean-cycle4 and HP-cycle4 in relation to CTE4. The Siegel-Tukey test ranks the series from smallest to largest value. It assigns rank 1 to lowest value, rank 2 to highest value, the rank 3 to the second highest value, and rank 4 to the second lowest value, rank 5 to the third lowest value, and so on. *In other words, the ranking for the Siegel-Tukey tests alternates from the lowest to the highest value for every other rank* (Eviews, p.163). As in Kruskal-Wallis test, the idea is to compare the sum of the ranks from each pair of subgroups (1, 2, …, n). If the groups have the same variances, the values of the summations of the respective ranks should be similar.

Furthermore, we proceed with the comparison of amplitude of the cycles as follows.

$VAR(CTE4) = (0.346131)^2$,

$VAR(DDME4) = (0.390962)^2$, and (18)

$VAR(HP4) = (0.346965)^2$.

Using these resulting to account for the amplitude of each estimate of CTE4, we have:

$$\left(\frac{0.390962}{0.346131}\right)^2 > \left(\frac{0.346965}{0.346131}\right)^2 \Rightarrow 133\% > 100\%. \quad (19)$$

Thus, in this model HP-cycle shows to be highly superior in relation to the cycle created by DDME, in the sense that it mimics closely the amplitude of the CTE4. However, we may show that the true cycle requires one step forward to be estimated. Before that, it is worth to compare one more result about the significance of each estimator of the cycle in the models. Table 1 displays the resulting explanation coefficients for each model.

Table 1
OLS Regressions on CTE and their respective statistics

Dependent Variable	Variable	Coefficient	Std. Error	t-Statistic	Probability	R^2
CTE1	Mean-cycle1	1.0	0.009	114.84	0.000	0.99
	HP-cycle1	1.0	0.022	46.89	0.000	0.96
CTE2	Mean-cycle2	1.1	0.012	91.93	0.000	0.99
	HP-cycle2	1.1	0.023	48.16	0.000	0.96
CTE3	Mean-cycle3	0.62	0.049	12.50	0.000	0.63
	HP-cycle3	0.70	0.053	13.17	0.000	0.66
CTE4	Mean-cycle4	0.77	0.045	17.25	0.000	0.77
	HP-cycle4	0.88	0.043	20.58	0.000	0.82

In accordance with Table 1, for models I and II, the cycle generated through MMDE procedure produced higher R-squared than those from the application of HP filter. In addition, models III and IV the R-squared of HP-cycles were higher than those R^2 resulting from the application of MMDE method.

V. The True Cycle Estimation

Recall that the CTE is the sum of the cycle (CL_{it}) and the irregular term (ER_t). In this section, we proceed with the comparison of the true cycle (CL_{it}) generated by the Monte Carlo experiment and the ones resulting from the use of HP-Filter and the DDME method of smoothing.

One intriguing result was the estimation of the cyclical component of Model I using the Hodrick-Prescott filter. Notice that HP trend smoothing does not separate the irregular term of the variable from its cyclical term. Thus, the estimated cycle using HP filter technique showed to be the opposite of the true value of the cyclical component created by the Monte Carlo experiment.

Additionally, when we applied the HP filter to estimate the cyclical components to models II and III, we confirmed the same opposite direction of the sign of the true cycle as we found in Model I. Therefore through the HP filtering it was unable to determine the true cyclical components of Y_2 and Y_3. We expected this to be the case because the HP filter considers that the irregular component of the series is smooth or insignificant. Especially for the case of the second model, the stochastic trend was relatively complex, making it difficult to determine the true cycle. Figures 13 through 16 display these findings.

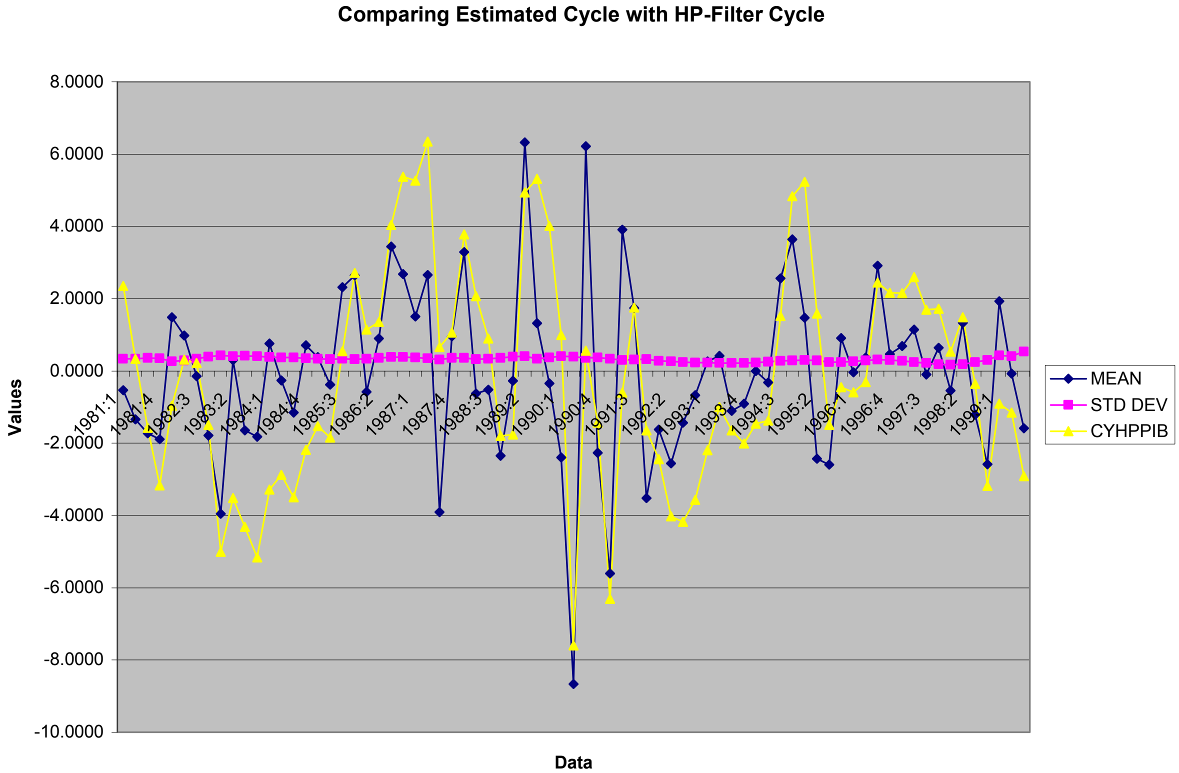
VI. An Application to Brazilian GDP

In line with the recursive procedure presented here, we applied this new technique to the Brazilian product (Gross Domestic Product-PIB) to the period of 1980:I to 1999:III. The data comes from IPEA (*Instituto de Pesquisa Econômica Aplicada*). The objective was to compare the resulting cyclical component of DDME technique with HP-Filter. Both calculations are represented in Graph 1.

Since we cannot specify the stochastic component of real Brazilian product, we estimated the mean-cycle for Brazilian GDP throughout our recursive technique and called it Mean-Cycle PIB to compare with the resulting cycle form HP filtering.

In accordance with Graph 1, the Mean-Cycle PIB mimics the HP-Cycle, they apparently follow the same path. However, it seems that the variations of the Mean-Cycle PIB are lower for the most part of the data. Its standard deviation series is shown in the smooth line close to the horizontal axes.

Graph 1



VII. Bibliography

Hansen, Bruce E. (2000). Testing for Structural Change in Conditional Models. *Journal of Econometrics,* vol. 97: 93-115.

Burns, A.F. & Mitchell, W.C., 1946. *Measuring Business Cycle.* National Bureau of Economic Research, New York, NY.

Cooley, T.F. (editor), 1995. *Frontiers of Business Cycle Research.* Princeton University Press, Princeton, NJ.

Hamilton, J.D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica,* Vol. 57, No. 2, March, pp. 357-384.

Harvey, A.C. & Jaeger, A., 1993. Detrending, Stilized Facts and the Business Cycle. *Journal of Applied Econometrics*, Vol. 8, No. 3, Jul-Sep., pp. 231-247.

Hodrick, R.J. & Prescott, E.C., 1980. Postwar U.S. Business Cycles: An Empirical Investigation. *Working Paper,* Carnegie-Mellon University, Pittsburgh, PA.

____________________________, 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking,* Vol. 29, No. 1, February, pp.1-16.

Maddala, G.S. & Kim, I-M. (1998). *Unit Roots, Cointegration, and Structural Change.* Cambrigde University Press, New York, NY. 505p.

Nelson, C.R. & Plosser, C.I., 1982. Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications. *Journal of Monetary Economics,* Vol. 10, pp.139-162.

Perron, P., 1988. Trends and Random Walks in Macroeconomic Time Series. Further Evidence from a New Approach. *Journal of Economic Dynamics and Control*, Vol. 12, pp. 297-332.

Perron, P., 1990. Testing for a Unit Root in a Time Series with a Changing Mean. *Journal of Business and Economic Statistics*, April, Vol. 8, Nº 2, pp. 153-162.

Perron, P. & Vogelsang,T.J., 1992. Nonstationarity and Level Shifts with an Application to Purchasing Power Parity. *Journal of Business and Economic Statistics*, July, Vol.10, Nº 3, pp. 301-320.
Phillips, P.C.B., 1987. Toward a Unified Asymptotic Theory for Autoregression. *Biometrika*, Vol.74, Nº 3, pp.535-547.
Phillips, P.C.B., & Hansen, B.E., 1990. Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies*, 57, pp. 99-125.
Phillips, P.C.B. & Perron, P., 1988. Testing for a Unit Root in Time Series Regression. *Biometrika,* Vol. 75, pp. 335-346.
Vogelsang, T.J., 1997. Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series. *Econometric Theory*, 13, pp. 818-849.
Zivot, E. & Phillips, P.C.B., 1994. A Bayesian Analysis of Trend Determination in Economic Time Series. *Econometric Reviews*, Vol.13, Nº 3, pp. 291-336.

VIII. Appendix A

The mathematical formulations of the general model can be represented as follows.

$$y_t = \theta_t + f(t) + \delta y_{t-1} + \sum_{i=2}^{n=t} y_{i-1}$$

Where y_t is a time series, θ_t is the parameter representing the intercept of the variable, and t is the trend of the variable. In accordance with this framework, θ_t follows the form bellow.

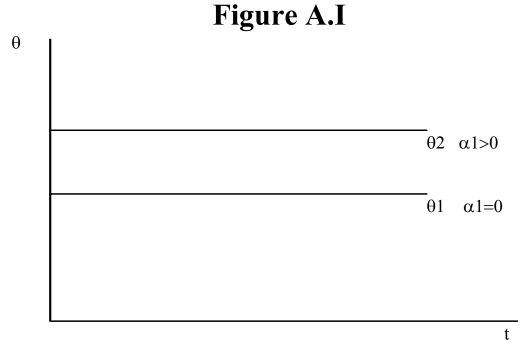
$$\theta_t = \alpha_0 + \alpha_1 DC_1$$

DC_1 can take more than one form, such as:

1) If $DC_1 = 0$, then $\theta_{1t} = \alpha_0$
2) If $DC_1 = 1$, then $\theta_{2t} = \alpha_0 + \alpha_1$.

Figure A.I displays these possibilities.

Figure A.I



$DC_1 = 0$ for $t < \bar{t}_i$ and $DC_1 = 1$ for $t \geq \bar{t}_i$, where i=1, …, n- $\bar{t}$. Understanding this presentation gets easy in the light of an example.

Example A.I: $t > \bar{t}_i$ and i=3

Obs. = t	DC_1	α_0
1	0	1
2	0	1
3	0	1
4	1	1
5	1	1
:	:	:
n	1	1

The function of time [f (t)] is the following:

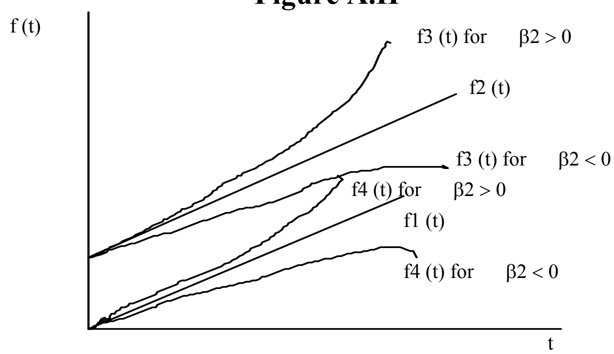
$$f(t) = \beta_0 + \beta_1 t + \beta_2 D_1 \bar{t} .$$

Where D_1 is a dummy variable to capture possible changes in t. Thus D_1 can assume different forms, such that:

1) If $D_1 = 0$ and $\beta_0 = 0$, then $f_1(t) = \beta_1 t$,
2) If $D_1 = 0$ and $\beta_0 \neq 0$, then $f_2(t) = \beta_0 + \beta_1 t$,
3) If $D_1 = 1$ and $\beta_0 \neq 0$, then $f_3(t) = \beta_0 + (\beta_1 t + \beta_2 \bar{t})$,
4) If $D_1 = 1$ and $\beta_0 = 0$, then $f_4(t) = (\beta_1 + \beta_2)\, t$.

These cases are drawn in Figure A.II.

Figure A.II



The qualitative variable $D_1 = 0$ for $t < \bar{t}_i$, $t = 1, 2, \ldots, n$. And $D_1 = 1$ for $t \geq \bar{t}_i$, where $i=1, \ldots, n-\bar{t}$. Considering one example in which $t \geq \bar{t}_i$ and i=3, and another one with $t \geq \bar{t}_i$ and i=2.

Example A.II: $t \geq \bar{t}_i$ and i=3

Obs. = t	D_1	$\bar{t}_i$
1	0	0
2	0	0
3	1	1
4	1	2
5	1	3
:	:	:
N	1	$n-\bar{t}=n-3$

Example A.III: $t \geq \bar{t}_i$ and i=2

Obs. = t	D_1	$\bar{t}_i$
1	0	0
2	1	1
3	1	2
4	1	3
5	1	4
:	:	:
n	1	n-1

Combining all the possibilities considered here for θ and f (t) forms a model to capture different features of intercept changes and stochastic trend of a series. In what follows we exhibit the resulting model we are interested into.

Considering the variable y_t.

$$y_t = \theta_t + f(t) + \delta y_{t-1} + \sum_{i=2}^{n=t} y_{i-1}$$

<u>Examples</u>:

Case I

DC=0, β_0 and D=0, then $y_t = \alpha_0 + \beta_1 t$;

Case II

DC=1, β_0 and D=0, then $y_t = \alpha_0 + \alpha_1 + \beta 1t$;

Case III

DC=0, $\beta_0 \neq 0$, and D=0, then $y_t = (\alpha_0 + \beta_0) + \beta_1 t$;

Case IV

DC=0, $\beta_0 \neq 0$, and D=1, then $y_t = (\alpha_0 + \beta_0) + \beta_1 t + \beta_2 \bar{t}$;

Case V

DC=1, $\beta_0 \neq 0$, and D=1, then $y_t = (\alpha_0 + \alpha_1 + \beta_0) + \beta_1 t + \beta_2 \bar{t}$;

Case VI

DC=1, $\beta_0 = 0$, and D=1, then $y_t = (\alpha_0 + \alpha_1) + \beta_1 t + \beta_2 \bar{t}$;

Case VII

DC=0, $\beta_0 = 0$, and D=1, then $y_t = \beta_1 t + \beta_2 \bar{t}$;

Case VIII

DC=0, $\beta_0 = 0$, and D=0, then $y_t = \beta_1 t$;

Case IX

DC=0, $\beta_0 = 0$, $\beta_1 = 0$, and D=0, then $y_t = \alpha_0$.

Notice that at $\bar{t}$, we could have:

1) θ_2, (α_1>0), and $f_1(t)$;
2) θ_2, ($\alpha_1 < \alpha < \alpha_1$'), and $f_4(t)$ with ($\beta_2 < 0$);
3) θ_2, ($\alpha_1 < \alpha < \alpha_1$”), and $f_4(t)$ with ($\beta_2 > 0$).

Figure A.IV illustrates some of the possible cases described here. Each point ($\bar{t}_i$) is estimated in order to obtain the best fit for y_t. The $\hat{y}_t$ and the residuals are then saved. At the end of running all regressions, what you have is an array of estimated $\hat{y}_t$ and residuals forming a matrix. The mean of the first line, with the estimated residuals from the regressions performed for that observation, represents the estimated cycle together with its error term.

Figure A.III

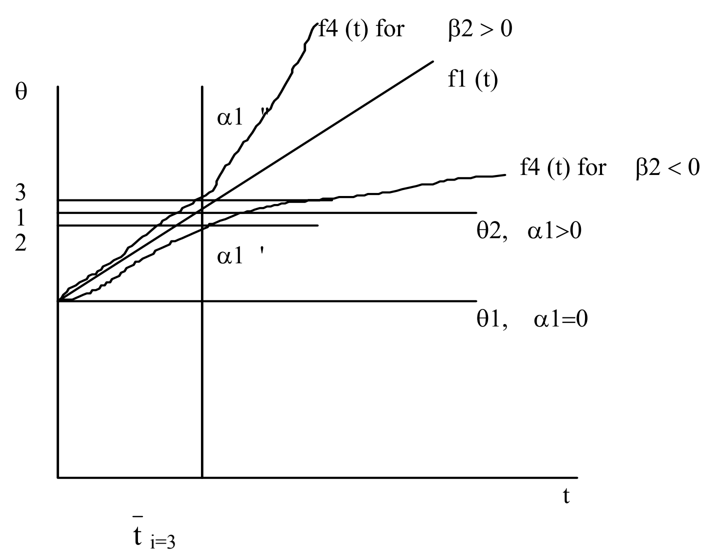


Figure 1

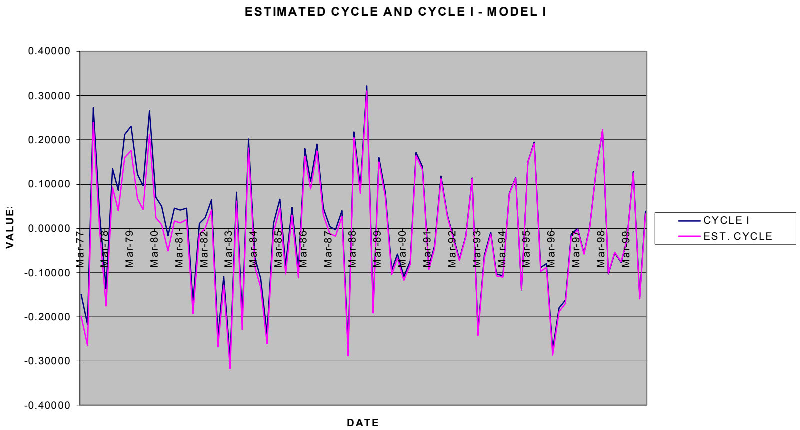


Figure 2

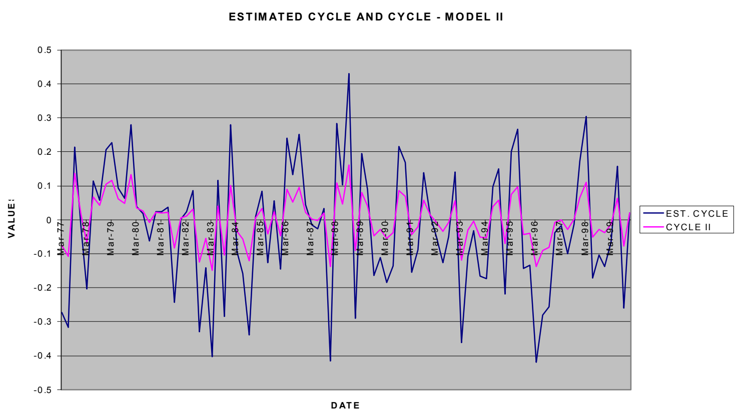


Figure 3

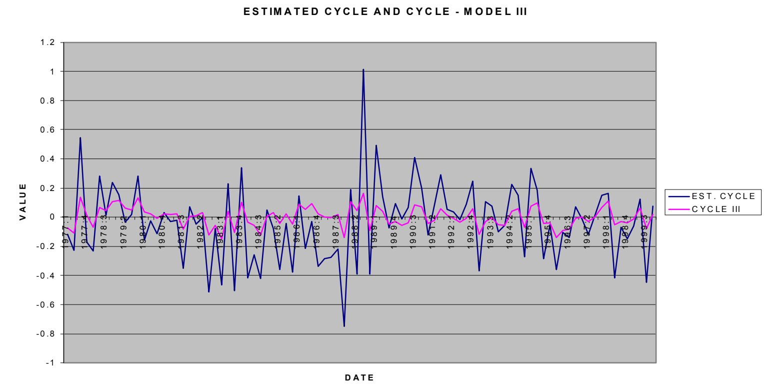


Figure 4

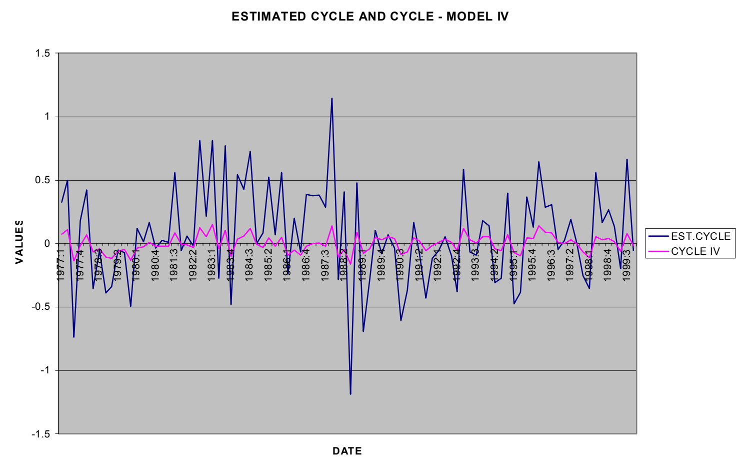


Figure 5

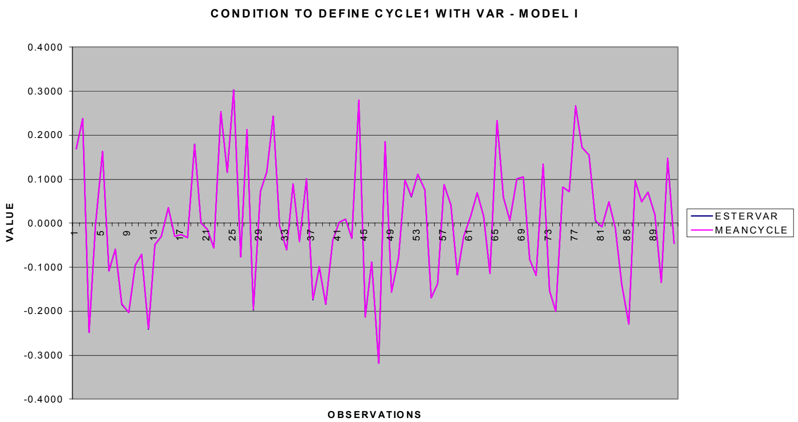


Figure 7

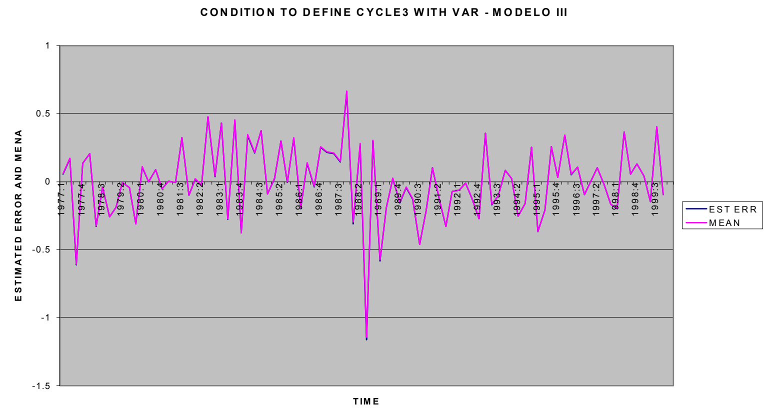


Figure 6

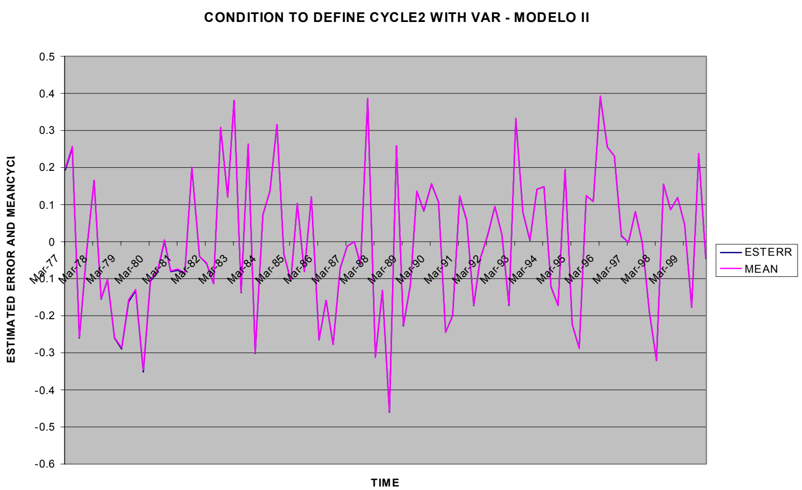


Figure 8

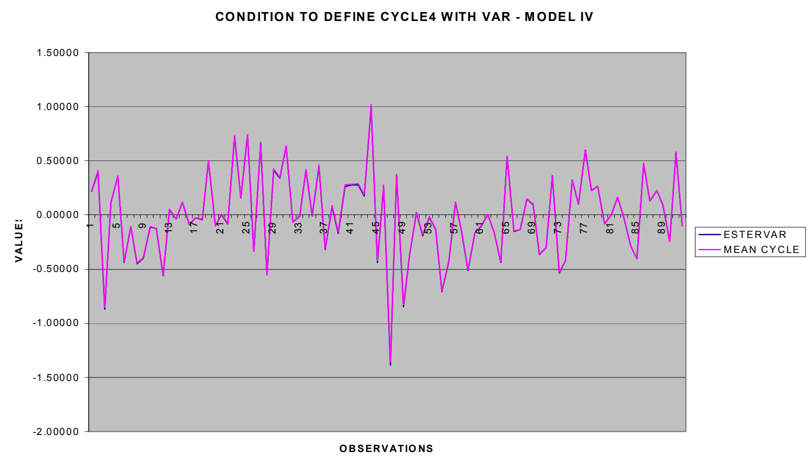


Figure 9

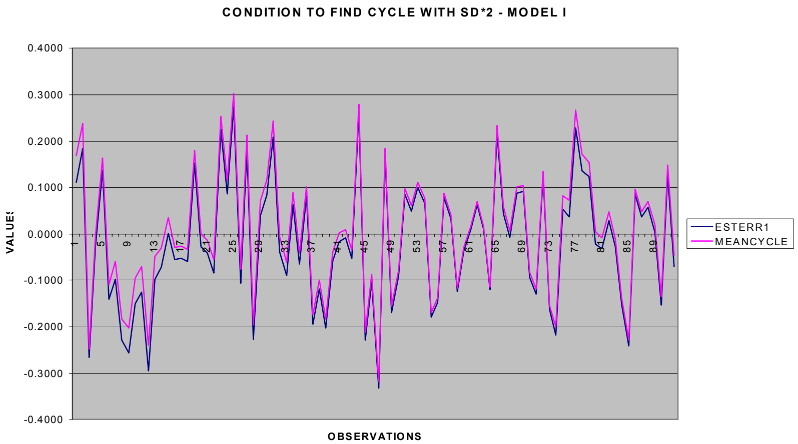


Figure 11

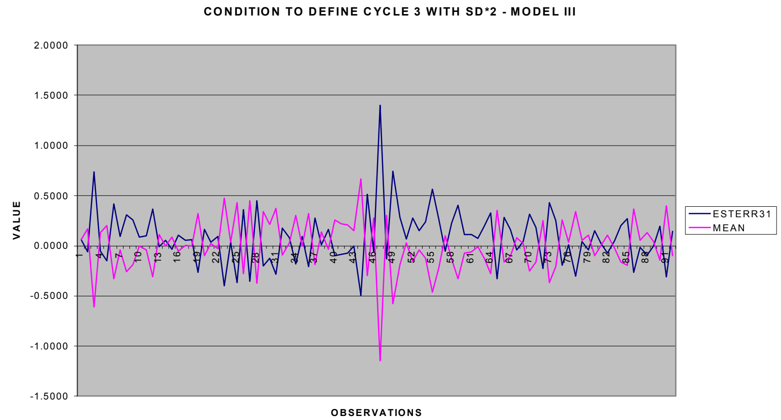


Figure 10

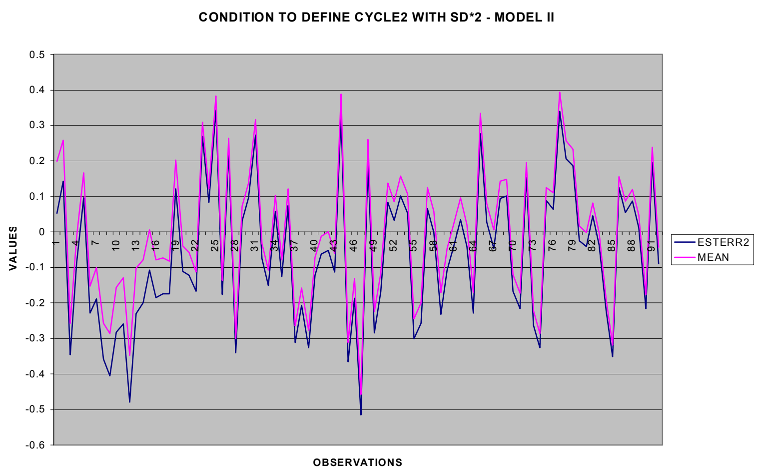


Figure 12

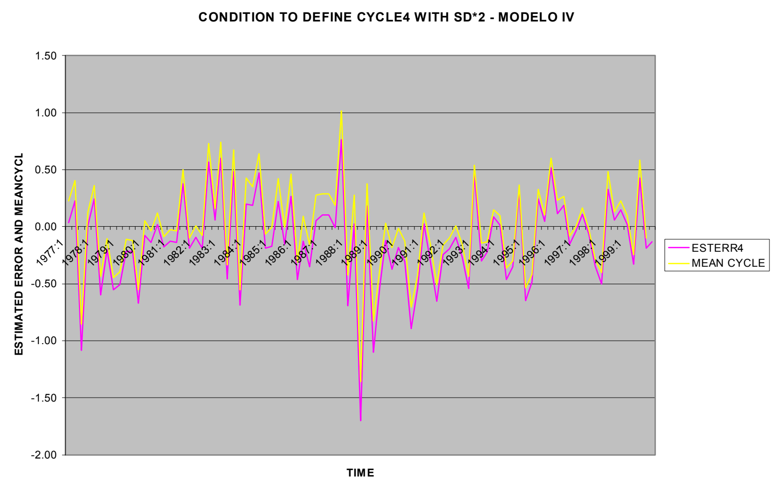


Figure 13

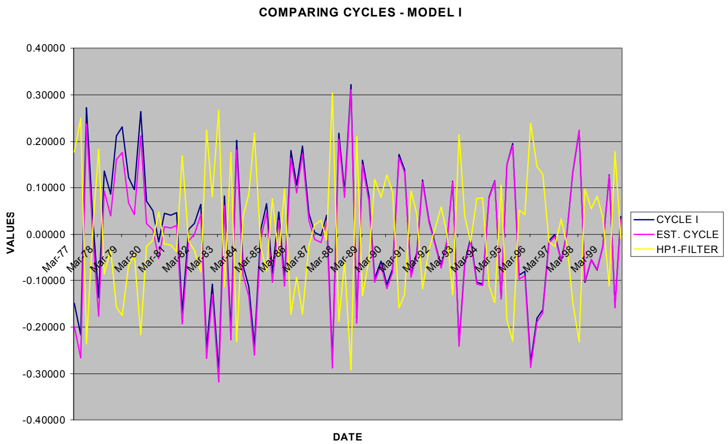


Figure 14

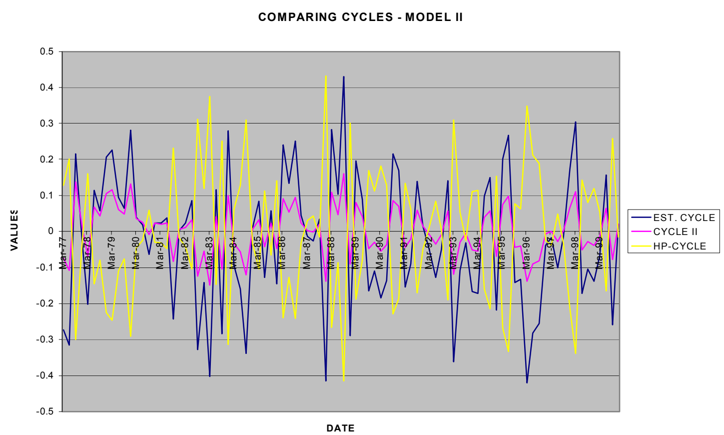


Figure 15

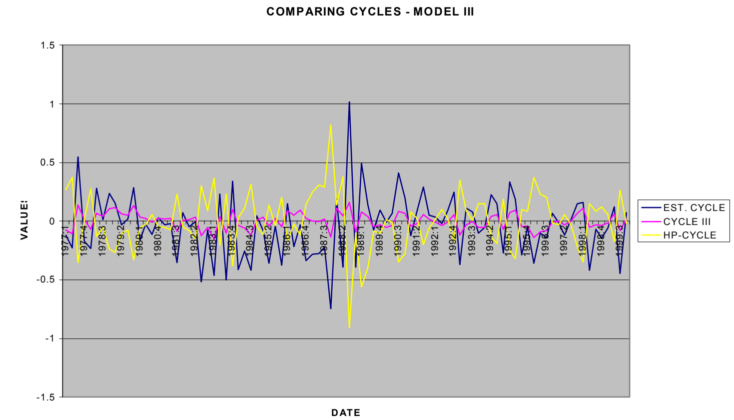


Figure 16

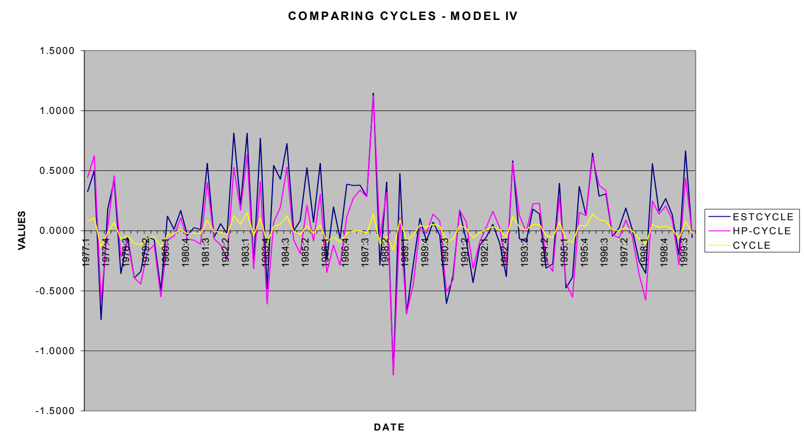


Figure 17

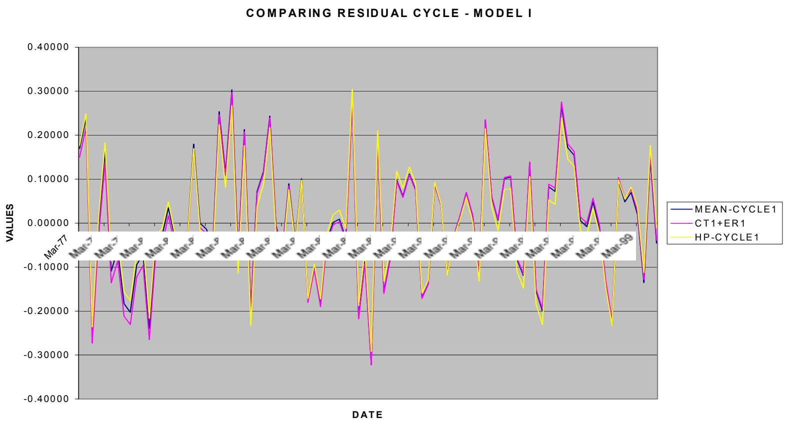


Figure 18

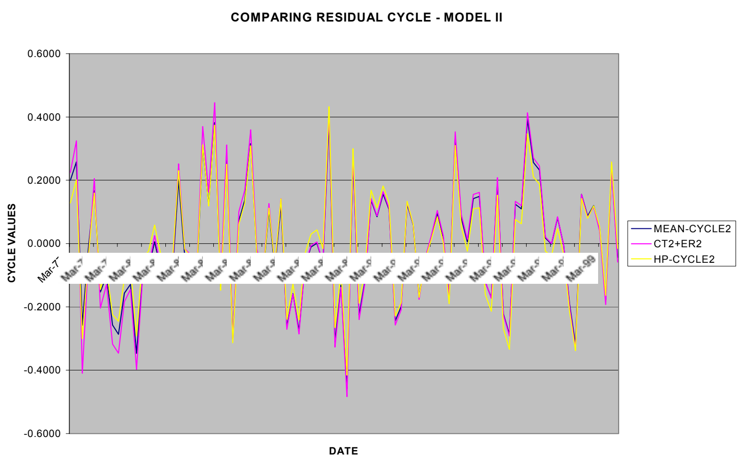


Figure 19

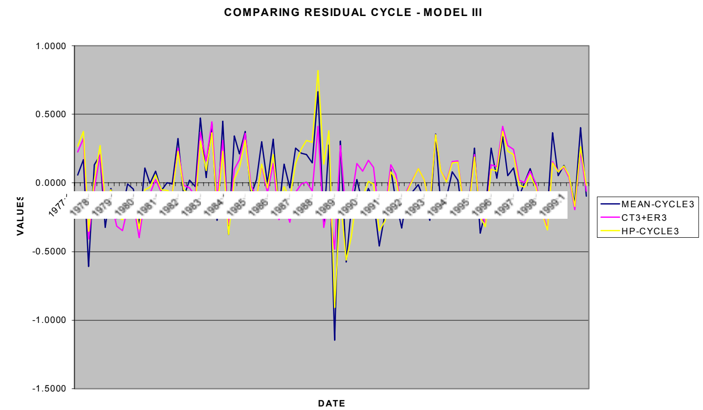


Figure 20