

Filtering Methods Revisited

Rafael Doménech^{*}, Víctor Gómez^{} and David Taguas^{**}**

D-97005

July, 1997

*** University of Valencia and Ministry of Economy and Finance.**

**** Ministry of Economy and Finance.**

The Working Papers of the Dirección General de Análisis y Programación Presupuestaria are not official statements of the Ministerio de Economía y Hacienda.

This paper was presented at the XXI Simposio UAB. The authors thank J. Fernández-Villaverde and J.E. Galdón for helpful comments. R. Doménech also thanks CICYT for the financial support given by SEC 96-1435.



Abstract

This paper focuses on the issue of trend and cycle decomposition. Firstly, three well-known univariate methods of stochastic detrending are reviewed and some new results concerning these methods are reported. The methods are the Hodrick-Prescott (HP) filter, the Beveridge-Nelson (BN) decomposition and a symmetric moving average filter proposed by Baxter and King (BK). Then, a two-step model-based procedure is proposed to circumvent some of the problems associated with *ad-hoc* filtering. The first step consists of using a model-based procedure to estimate the trend-cycle signal, whereas the second step applies a fixed low-pass filter to this estimate to obtain the trend and cyclical components. The proposed procedure presents four advantages over the alternative methods. First, the trend-cycle component obtained in the first step depends on the characteristics of the series. Second, the series is corrected for outlier and other deterministic effects. Third, the fixed filter for annual data used in the second step approximates the gain of the ideal filter better than the HP filter and finally it requires fewer parameters than the BK filter.



1. Introduction

Although filters have been widely used in economics, it was not until the seminal work of Kydland and Prescott (1982) that they began to play the central role they have nowadays. The main reason behind this phenomenon is that, in the Real Business Cycle Literature, the usual way to evaluate calibrated models is to compare some statistics of the simulated series with those of the filtered observed series. Recently, a lot of work concerning the improvement of the behaviour of some of those statistics has been done. Some examples are the standard deviation of investment or the correlation between productivity and hours worked (for a general review, see Hansen and Wright (1994)), or with important theoretical advances like the use of heterogeneous agents. However, much less thought has been given to the possibility of improving the statistical tools used to filter the series. Since the work of Kydland and Prescott (1982), the standard approach for filtering economic time series has been the use of the Hodrick-Prescott filter (1997), henceforth HP.

Despite the fact that a relatively common element of this extensive literature has been that the empirical measurement is based on quarterly US data, in the last few years, the growing popularity of these models has motivated numerous applications to other countries, mainly in the area of the OECD. Some of these countries have not quarterly data available or the length of the sample period is too short to compute adequate statistics. As a consequence, there is a growing literature that uses annual data to evaluate simulated series. For example, both Correia, Neves and Rebelo (1995), with regard to the Portuguese economy, and Backus and Kehoe (1992), in their study of international business cycles, use annual data with a smoothing parameter of the HP filter equal to 100. Over and above, besides the traditional use of the HP filter in the Real Business Cycle Literature, in the last few years alternative applications of this filter have arisen, both at international and national organizations. Thus, it is very common the use of the HP filter to estimate the trend component of the GDP, with the purpose of finding something resembling the potential output, or to compute structural and cyclical budget deficits, using annual data. In all these cases, these exercises are also carried out with a smoothing parameter which is equal to or higher

than 100.

Although the HP filter with a smoothing parameter equal to 1600 often performs rather well for quarterly data, this paper shows that the use of this filter with annual data and values of the smoothing parameter as high as 400, or even the substantially lower 100, can be misleading and proves that a smoothing parameter of 10 is more adequate in this case. This result, which is obtained by analyzing the gain function of the filter, coincides with the one reported by Baxter and King (1995).

At the same time, some authors have raised further concerns about the qualities of the HP filter (as, for example, King and Rebelo (1993), Harvey and Jaeger (1993) and Cogley and Nason (1995)) and others have proposed alternative procedures like Baxter and King (1995) or, previously, Beveridge and Nelson (1981). All the above reasons are enough to think about filters from an alternative perspective rooted in the tradition of Butterworth filters commonly used by engineers. Thus, three different univariate methods of stochastic detrending are reviewed and several new results concerning these methods are presented. These methods are the HP filter, the Beveridge-Nelson decomposition, henceforth BN, and the symmetric moving average filter proposed by Baxter and King (1995), henceforth BK.

After reviewing these univariate methods, a new two-step procedure to estimate the trend and cycle components is proposed. The first step consists of using a model-based signal extraction procedure to estimate the trend-cycle component, whereas the second step applies a fixed low-pass filter to this trend-cycle to obtain the trend.¹ The cycle is computed as the difference between the first and the second estimated components. The idea behind this method is to overcome some of the problems associated with *ad-hoc* filtering.

The method proposed in this paper displays four advantages over the alternative filters. First, it is possible to prevent a bad use of the filter by analyzing the original series and accounting for its characteristics of the series. Second, the series can be corrected for outliers, which can significantly alter the

¹ This procedure can be implemented using the programs TRAMO and SEATS developed by Gómez and Maravall (1997).

estimation of the trend and cycle components. Third, the proposed fixed filter is a better approximation than the HP or the BN filter to the ideal filter, specially for annual data. Fourth, it needs fewer parameters and predictions than the BK filter.

The structure of this paper is as follows. In the second section we review three different filtering methods to obtain the business cycle and some aspects of spectral theory associated with them, which will be useful to interpret their performance when applied to usual time series. We also analyze the main characteristics of our proposed filter. In the third section we compare the performance of this fixed filter with that of the HP filter, applying both methods to the GDP of OECD countries from 1955 to 1995, using annual data. Different exercises show how the use of these filters alters some usual statistics such as volatility or international comovements of cyclical components across countries. Finally, the paper concludes with the main results in the fourth section.

2. Different Filtering Methods to Obtain the Business Cycle

2.1 Linear Time Invariant Filters and the Business Cycle: some analytical tools

Given an infinite time series $\{z_t\}$, a linear time invariant filter that produces a new time series $\{\bar{z}_t\}$ is a transformation of the form

$$\bar{z}_t = \sum_{j=-\infty}^{\infty} h_j z_{t-j}$$

In terms of the lag operator, defined by $Lz_t = z_{t-1}$, the filter can be expressed more compactly as

$$\bar{z}_t = H(L)z_t = \sum_{j=-\infty}^{\infty} h_j L^j z_t, \quad (1)$$

where $H(L) = \sum_{j=-\infty}^{\infty} h_j L^j$ and $L^j = L(L^{j-1})$. The filter is said to be symmetric if $h_{-j} = h_j$ for all $j = 1, 2, \dots$, in which case $H(L)$ can be written as

$$H(L) = h_0 + \sum_{j=1}^{\infty} h_j (L^j + L^{-j})$$

Every stationary process $\{z_t\}$ has a representation in terms of periodic stochastic components that resembles the representation of a periodic function as a Fourier series. This is the so-called spectral representation, which mathematically is expressed as

$$z_t = \int_{-\pi}^{\pi} e^{itx} dY(x) \quad (2)$$

Here, $\{Y(x)\}$ is a stochastic process with independent increments, continuous on the right, and $i = \sqrt{-1}$. Heuristically, the previous integral can be

interpreted as the limit in mean squared of sums of sinusoids with stochastic coefficients. The effect of the linear filter (1) in the spectral representation (2) is the following. Define the function $\hat{H}(x)$ as the Fourier transform of the sequence $(\dots, h_{-1}, h_0, h_1, \dots)$, that is

$$\hat{H}(x) = \sum_{j=-\infty}^{\infty} e^{-ijx} h_j, \quad (3)$$

which is obtained by replacing L in $H(L)$ with e^{-ix} . Then, it can be shown that the spectral representation of the transformed series \bar{z}_t is

$$\bar{z}_t = \int_{-\pi}^{\pi} \hat{H}(x) e^{itx} dY(x) \quad (4)$$

The function $\hat{H}(x)$ is usually called the frequency response function of the filter. Equation (4) shows intuitively the effect of the filter $H(L)$ on the input series $\{z_t\}$. The effect on the sinusoids at frequency $x \in [-\pi, \pi]$ is twofold. First, the amplitudes are multiplied by the modulus $G(x) = |\hat{H}(x)|$ of the complex number $\hat{H}(x)$ and second, there is a shift effect measured by the argument $\phi(x)$ of $\hat{H}(x)$. The functions $G(x)$ and $\phi(x)$ are usually called gain and phase function of the filter, respectively. Note that if the filter is symmetric, there is no phase effect, since the number $\hat{H}(x)$ is a real number due to the cancellation of the sine functions in the expression (3). This feature makes symmetric filters desirable in all practical applications.

Given the autocovariances $\gamma(j) = E(z_{t+j}z_t)$ of a stationary process $\{z_t\}$, which is assumed to follow an *ARMA* model, the spectrum of the process $f(x)$ is defined as the Fourier transform of the covariance sequence $(\dots, \gamma(-1), \gamma(0), \gamma(1), \dots)$, that is

$$f(x) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-ijx}$$

For example, for a white noise series $\{a_t\}$ with $Var(a_t) = \sigma_a^2$, the spectrum is $f(x) = \sigma_a^2/2\pi$. The autocovariances $\gamma(j)$ are given by the inverse Fourier

transform of $f(x)$

$$\gamma(j) = \int_{-\pi}^{\pi} e^{itx} f(x) \quad (5)$$

If $f_z(x)$ is the spectrum of the series $\{z_t\}$ in (1), then the spectrum $f_{\bar{z}}(x)$ of the filtered series $\{\bar{z}_t\}$ is given by the formula

$$f_{\bar{z}}(x) = |\hat{H}(x)|^2 f_z(x), \quad x \in [0, \pi]$$

As an application of this formula, suppose that the series $\{z_t\}$ follows the *ARMA* model $\phi(L)z_t = \theta(L)a_t$, where $\text{Var}(a_t) = \sigma_a^2$. Then, the series $\{z_t\}$ can be seen as the result of applying the filter $H(L) = \theta(L)/\phi(L)$ to the white noise series $\{a_t\}$. Therefore, the spectrum is

$$f(x) = \frac{1}{2\pi} \frac{\theta(e^{-ix})\theta(e^{ix})}{\phi(e^{-ix})\phi(e^{ix})} \sigma_a^2 = \frac{1}{2\pi} \frac{|\theta(e^{-ix})|^2}{|\phi(e^{-ix})|^2} \sigma_a^2$$

Note that if $j = 0$ en (5), we obtain the variance of the process expressed as an integral of the spectrum. This can be seen as how the stochastic components associated with the different frequencies contribute to the variance of the series.

Traditionally, filters have been designed in terms of their desired effect on a certain band of frequencies. So, a low-pass filter is a filter which, ideally, has a gain function equal to one for the frequencies near to zero and a value of zero for all the other frequencies. More generally, a band-pass filter for the band $[x_1, x_2]$ is a filter such that the gain function $G(x) = 1$ for $x \in [x_1, x_2]$ and is zero otherwise. Note that frequency and period are related by the well-known formula $x = 2\pi/T$, where x is the frequency and T is the period.

Before describing the different techniques for measuring business cycles, we have to address a central issue, namely, the *definition* of a business cycle. Nowadays it is widely accepted that business cycles are cyclical components of no more than thirty two quarters (eight years), and no less than six quarters in duration (Prescott (1986) and Baxter and King (1995)). If this definition is accepted without criticism, the ideal filter to obtain the business cycle would be a band-pass filter which passes all frequencies in the desired

band. In order to obtain the ideal filter for yearly data, one has to take into account the fact that the greatest detectable period is two. Therefore, the ideal filter passes all stochastic components with frequencies greater than or equal to $2\pi/8$. This filter can be seen in Figure 1.

2.2 The Hodrick-Prescott Filter

In recent years, there has been a dramatic growth in the use of the filter proposed by Hodrick and Prescott (1980), although the idea that underlies the development of this filter seems to be an old one. Thus, Gersch and Kitagawa (1990) trace back this idea to Whittaker (1923). Following Gersch and Kitagawa (1990), the problem treated by Whittaker consisted of decomposing the given observations z_t , $t = 1, \dots, N$, as the sum of a "smooth" function \bar{z}_t and observation noise

$$z_t = \bar{z}_t + z_t^c \quad (6)$$

Here, one tries to estimate the unknown \bar{z}_t , $t = 1, \dots, N$, which in a typical application is the trend of a nonstationary time series. Whittaker suggested that the solution balance a trade-off of goodness of fit to the data and goodness of fit to a smoothness criterion. This idea was realized by minimizing

$$\sum_{t=1}^N (z_t - \bar{z}_t)^2 + \lambda \sum_{t=k+1}^N (\nabla^k \bar{z}_t)^2, \quad (7)$$

where $\nabla^k = \nabla(\nabla^{k-1})$, for some appropriately chosen smoothness trade-off parameter λ .

The properties of the solution to the problem (6)-(7) are clear. If $\lambda = 0$, $\bar{z}_t = z_t$ and the solution is a replica of the observations. As λ becomes increasingly large, the smoothness constraint dominates the solution and this satisfies a k th-order constraint. For large λ and $k = 1$, the solution is a constant; for $k = 2$, it is a straight line; and so on. Whittaker left the choice of λ to the investigator. The notable feature is the success of Hodrick and Prescott in choosing for quarterly data of the US economy a value of $\lambda = 1600$ and $k = 2$ in (7). We will return to this point later in this subsection.

To obtain the HP filter, define the $(N - 2 \times N)$ matrix A

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & -2 & 1 \end{pmatrix}$$

The minimization problem (7) can be reformulated as that of minimizing with respect to $\bar{z} = (\bar{z}_1, \dots, \bar{z}_N)'$ the function $(z^c)'z^c + \lambda(A\bar{z})'(A\bar{z})$, where $z^c = (z_1^c, \dots, z_N^c)'$. Using standard matrix differentiation results, the solution can be easily seen to be

$$\bar{z} = (I + \lambda A'A)^{-1}z,$$

where $z = \bar{z} + z^c$ is the observed series. The estimate of the business cycle is then

$$z^c = (I - (I + \lambda A'A)^{-1})z \quad (8)$$

The filter characteristics are best seen in the frequency domain. Since it is a finite filter, the filtered series is of the form $z_t^c = \sum_{h=1}^N d_{ht}z_h$, where the d_{ht} are the weights given by (8). In terms of the lag operator L , the filter can be expressed as $d_t(L) = \sum_{h=1}^N d_{ht}L^{t-h}$ and the filtered series as $z_t^c = d_t(L)z_t$. It is not a time-invariant filter and, therefore, the weights depend on the date t as well as the lead/lag index h . For each date t , the frequency response function of the filter is $\hat{H}_t(x) = d_t(e^{-ix}) = \sum_{h=1}^N d_{ht}e^{i(h-t)x}$, from which the characteristics of the filter can be obtained. As shown in Baxter and King (1995), the filter can behave very differently from the ideal filter at both ends of the series, which means that the estimates of z_t^c for dates t close to both ends of the series can be very poor.

This fact should be kept in mind by those who believe that the HP filter does not involve any cost in terms of forecasts or backcasts. The situation is precisely the opposite, that is, in order to improve the estimates at both ends of the series, it is desirable to extend the series with forecasts and backcasts obtained with some model, typically an *ARIMA* model. We will return to

this point later. This property is also shared by all symmetric moving average filters.

If we had models for the components in (6), it would be possible to have a better understanding of the filter properties. In addition, by applying the Wiener-Kolmogorov filters, we would obtain the infinite sample version of the HP filter. Thus, the finite sample HP filter would be an approximation to the infinite HP filter. The frequency response function of the infinite filter would be very useful to determine how close it is to the ideal filter.

It is a remarkable fact that the problem of estimating the trend \bar{z}_t in (6) by means of the minimization problem (7) is equivalent to estimating the signal \bar{z}_t in model (6), under the assumption that the components follow certain models. Specifically, \bar{z}_t is assumed to follow the model

$$\nabla^2 \bar{z}_t = \eta_t,$$

where the initial values \bar{z}_{-1} , \bar{z}_0 are taken as unknown constants or equivalently diffuse, the η_t is a sequence of i.i.d $N(0, \sigma_\eta^2)$ and the z_t^c is a sequence of i.i.d. $N(0, \sigma_\epsilon^2)$ independent of the η_t . A proof of this result can be seen in Kohn and Ansley (1988).

Using the models for the components, it is possible to derive the estimator of z_t^c , based on the doubly-infinite series $\{z_t\}$, which is supposed to be known. The estimator is given by the Wiener-Kolmogorov filter applied to the series, since, as shown by Bell (1984), the Wiener-Kolmogorov filter can also be applied to nonstationary time series. It is not difficult to show that the formula for the estimator is

$$\begin{aligned} \hat{z}_t^c &= \sigma_\epsilon^2 \frac{(1-L)^2(1-L^{-1})^2}{\sigma_\eta^2 + (1-L)^2(1-L^{-1})^2\sigma_\epsilon^2} z_t \\ &= \frac{\lambda(1-L)^2(1-L^{-1})^2}{1 + \lambda(1-L)^2(1-L^{-1})^2} z_t, \end{aligned}$$

where L is the lag operator and $\lambda = \sigma_\epsilon^2/\sigma_\eta^2$. Note that, since this filter is

symmetric, no phase shift is introduced. Since $z_t^c = z_t - \bar{z}_t$,

$$\hat{z}_t = \frac{1}{1 + \lambda(1 - L)^2(1 - L^{-1})^2} z_t$$

is the trend estimator.

The frequency response function $\hat{H}(x)$ of the infinite HP filter is obtained by replacing L with e^{-ix} in the previous expression for \hat{z}_t . After some manipulation, we have

$$\hat{H}(x) = \frac{4\lambda(1 - \cos(x))^2}{1 + 4\lambda(1 - \cos(x))^2}$$

Since this function is real, it coincides with the gain function of the filter. Further manipulation shows that

$$\hat{H}(x) = 1 - \frac{1}{1 + 16\lambda \sin^4(x/2)},$$

where the term after the minus sign in the previous formula is the gain function of the trend filter. This gain function coincides with the squared gain function of a Butterworth filter of the sine version, which has the general expression

$$|G(x)|^2 = \frac{1}{1 + \left[\frac{\sin(x/2)}{\sin(x_c/2)} \right]^{2N}} \quad (9)$$

These filters are low-pass filters that, ideally, pass all components with frequencies less than or equal to a frequency $x_l > x_c$, with squared gain equal to 1/2 for $x = x_c$. By choosing the parameters N and x_c , they can be designed to be as close as desired to the ideal filter.

Thus, for the HP filter, $N = 2$ and $\sin(x_c/2) = 1/(2\lambda^{1/4})$. In the case of quarterly data, $x_l = 2\pi/32$ and $x_c = 2\pi/39.7166$ if $\lambda = 1600$. The trend filter, and the HP filter too, has gain equal to 1/2 for the frequency x_c . For example, the quarterly HP filter has gain equal to 1/2 for the period of 39.7166 quarters. These facts are very important because they give an indication on how to choose the value of λ in the HP filter.

Butterworth filters of the sine version are autoregressive filters, with the degree of the autoregressive polynomial equal to N , which induce a phase effect because they are not symmetric. That is, the Butterworth filter $H_B(L)$ corresponding to (9) is of the form

$$H_B(L) = \frac{1}{\beta(L)},$$

where $\beta(L) = b_0 + b_1L + \dots + b_NL^N$ and $|G(x)|^2 = H_B(e^{-ix})H_B(e^{ix})$ is the squared gain of the filter. We can say that the trend filter $H_p(L)$ corresponding to the HP filter is a "symmetrized" form of the Butterworth filter, since it is of the form $H_p(L) = H_B(L)H_B(L^{-1})$.

Analyzing the gain function of the infinite HP filter for the different values of λ , the HP filter seems to approximate very well the ideal high-pass filter that corresponds to the frequency $x = 2\pi/32$ when $\lambda = 1600$, passing only those stochastic components with a period smaller than eight years. Notice that, when used with quarterly data in order to obtain the business cycle component, which corresponds to shocks with a period smaller than eight, the HP filter should be applied to seasonally adjusted data. This is so because the HP filter passes the seasonal component, which has peaks in the spectrum for the seasonal frequencies $2\pi/4$ and $4\pi/4$.

The situation is different when the HP filter is applied to data sampled yearly. In this case, the cut-off frequency is $2\pi/8$. This implies a value of λ considerably smaller than 400, frequently used in the literature. Applying (9), it can be seen that a value $\lambda = 10$ is more appropriate for annual data.²

However, the HP filter that corresponds to this value is not entirely satisfactory because it contains significant *leakage* as well as significant *compression*, as can be seen in Figure 2. In the filter jargon, *leakage* means that stochastic components with undesirable frequencies are passed and *compression* that some of the passed components with the correct frequencies are damped.

As mentioned above, in the more realistic situation, when only a finite series

² This value coincides with the one proposed by Baxter and King (1995) for annual data.

z_t , $t = 1, \dots, N$, is known instead of the whole realization, it should be noted that, although the HP filter gives estimates of the business cycle for all $t = 1, \dots, N$, the estimates given at both ends of the series can be very poor and it may prove useful to extend the series with forecasts and backcasts. This idea is similar to the one used in the X11-ARIMA program, which has proved useful in reducing revisions.

2.3 The Baxter and King Filter

Baxter and King (1995) propose a moving average band-pass filter to estimate the business cycle. The filter is a symmetric moving average. That is, for a series $\{z_t\}$, the filtered series $\{\bar{z}_t\}$ is

$$\bar{z}_t = \sum_{k=-K}^K a_k z_{t-k},$$

where $a_{-k} = a_k$, $k = 1, \dots, K$. The filter can be written more compactly using the lag operator L as $\bar{z}_t = a(L)z_t$, where $a(L) = \sum_{k=-K}^K a_k L^k$. These weights are chosen according to some optimality criterion and the band-pass filter is constructed as the difference between two high-pass filters.³

An advantage of the Baxter and King (BK) filter is that, being a band-pass filter, it does not pass frequencies higher than a preselected value, for example, $2\pi/6$ for quarterly data. Therefore, the BK filter can be applied to the original data and not necessarily to the seasonally adjusted data. However, like all moving average filters, the BK filter has the disadvantage that it cannot be applied at both ends of the data unless one can extend the series with forecasts and backcasts. This is a serious problem and, in this context, it is worthwhile to mention that there are autoregressive band-pass filters, like those of the Butterworth family, that can be as exact as the moving average filters, while requiring many less parameters and many less predictions.

³ See Baxter and King (1995) for the graphs of the gain function of their band-pass filter for different band widths.

2.4 *Ad-Hoc* Filtering: Dangers and Limitations

To illustrate the dangers of *ad-hoc* filtering,⁴ suppose that the HP filter is applied to a white noise series with unit variance (expressed, for convenience, in units of 2π) sampled quarterly. In this case, the spectrum has as a very wide peak for $x = \pi$, and hence the detrended series will behave as a stochastic component with a period of 2 quarters. Yet, by construction, the detrended series should be the white noise input.⁵

Another problem with *ad-hoc* filtering is the definition of the components. At first, it may seem that, by defining the component as the output of the filter (see Prescott (1986)), the issue of defining the component properly is simplified. This simplification is, however, misleading. This can be seen by considering an example. Suppose that the HP filter is applied to the last 35 years of quarterly US GNP. Assume we apply the HP filter to the first 18 years of the sample. According to the definition of Prescott, this estimator is the trend for $t = 72$. But if one more quarter is observed, the HP filter yields a different estimator of $t = 72$. Additional quarters will further change the estimator, and the HP filter is in fact a filter that implies a very long revision period. As Maravall (1993) has shown, the estimator fluctuates considerably and takes nearly 10 years to converge.

The lack of proper models for the components limits in many important ways the usefulness of *ad-hoc* filters. First, it makes difficult to detect the cases in which the filter is not appropriate for the series under study. Moreover, if such is the case, there is no systematic procedure to overcome the filter inadequacies. Second, even when appropriate, *ad-hoc* filtering does not provide the basis for rigorous inference. Given that the filter yields an estimator

⁴ Harvey and Jaeger (1993) illustrate with some examples the inadequate use of the HP filter

⁵ See Maravall (1993) for examples of misuse of the X11 filter when it is applied to quarterly series $\{z_t\}$ which follow the airline model

$$\nabla\nabla_4 z_t = (1 + \theta_1 L)(1 + \theta_4 L^4) a_t,$$

where $\{a_t\}$ is white noise, $\nabla_4 = 1 - L^4$ and the two moving average parameters lie between -1 and 1 .

of the unobserved component, it would be desirable to know the properties of the estimator, and, in particular, the underlying estimation errors. Further, *ad-hoc* filters do not provide the basis for obtaining forecasts of the components, which can also be of interest.

2.5 The Beveridge-Nelson Decomposition

Beveridge and Nelson (1981) have proposed a decomposition of a series $\{z_t\}$ into a permanent and a transitory component that allows for correlation between the two components. Supposing that $\{z_t\}$ is an $I(1)$ series such that ∇z_t has the Wold decomposition $\nabla z_t = \Psi(L)a_t$, then z_t can always be expressed as the sum of a permanent and a transitory component, where the permanent component is given by $\nabla \bar{z}_t = \Psi(1)a_t$, and the transitory component is equal to $c_t = \Psi^*(L)a_t$, where $\Psi^*(L)$ satisfies $(1 - L)\Psi^*(L) = \Psi(L) - \Psi(1)$. The Beveridge-Nelson decomposition can be seen as an ingenious decomposition of an $I(1)$ variable, but it does not properly fit into the unobserved components framework, since the components are, in fact, observable. This can be easily seen by rewriting, for example, \bar{z}_t as $\bar{z}_t = \Psi(1)\Psi(L)^{-1}z_t$, and hence both components are defined as linear combinations of the observed series.

A very easy form to obtain the Beveridge-Nelson decomposition is the following. Suppose that the series $\{z_t\}$ follows the *ARIMA* model $\phi(L)\nabla z_t = \theta(L)a_t$, where the polynomial $\phi(L)$ has all its roots outside the unit circle, and obtain the partial fraction expansion

$$\frac{\theta(L)}{\phi(L)\nabla} = \frac{k}{\nabla} + \frac{\theta_1(L)}{\phi(L)},$$

where k is a constant and $\theta_1(L)$ is a polynomial of degree less than or equal to that of $\theta(L) - 1$. Then, $\nabla \bar{z}_t = ka_t$ and $\phi(L)c_t = \theta_1(L)a_t$. Given that $a_t = [\phi(L)\nabla/\theta(L)]z_t$, the components can be expressed linearly in terms of the series $\{z_t\}$ as $\bar{z}_t = k[\phi(L)/\theta(L)]z_t$ and $c_t = [\nabla\theta_1(L)/\theta(L)]z_t$.

To illustrate, suppose we are interested in computing the cycle of the annual GDP for the United Kingdom $\{z_t\}$. Using program TRAMO the following

model was identified and estimated

$$\nabla z_t = 0.027 + (1 + 0.34L)a_t \quad (10)$$

Then, it is not difficult to check that

$$\frac{1 + 0.34L}{\nabla} = -0.34 + \frac{1 + 0.34}{\nabla}.$$

From this, it is obtained that $c_t = [-0.34/(1 + 0.34L)](\nabla z_t - 0.027)$. The gain function of this filter can be seen in Figure 4, from which it is deduced that the performance of the filter is rather poor. Since this series is finite, in order to apply the BN filter we need some starting conditions for the recursion $c_t = -0.34(c_{t-1} + z_t - z_{t-1} - 0.027)$. Using Tunnicliffe-Wilson's algorithm like in Burman (1980), it is obtained that $c_0 = -\frac{0.34}{1+0.34}(z_0 - z_{-1} - 0.027)$. The two backcasts are obtained by reversing the series and using model (10).

2.6 The Use of Fixed Filters within a Model-Based Approach

Using the automatic model identification procedure of TRAMO,⁶ we obtained the following model for the annual US GDP from 1950 to 1995:

$$\nabla z_t = 0.03 + a_t \quad (11)$$

The canonical model-based decomposition given by SEATS for this model is

$$z_t = p_t + u_t,$$

⁶ TRAMO and SEATS are two programs written in Fortran, developed by Gómez and Maravall (1997). They fall into the class of the so-called ARIMA-Model-Based methods for decomposing a time series into its unobserved components (i.e., for extracting from a time series its different signals). The components considered in SEATS are the trend, seasonal, cyclical and irregular components. Since the model that SEATS assumes is that of an integrated linear time series with Gaussian innovations, it is usually necessary to preadjust the series to be treated with SEATS with the program TRAMO. TRAMO removes from the series special effects, automatically identifies and removes several types of outliers, and interpolates missing observations. TRAMO can also be used as a standalone program for the detailed analysis of time series, since it performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of possibly several types of outliers.

where p_t is the trend component, which follows the model $\nabla p_t = (1 + L)b_t$, and b_t and u_t are uncorrelated white noise processes. For this series there are no peaks in the spectrum other than the big peak at the zero frequency, which is the trend frequency. The minimum mean squared error estimators \hat{p}_t and \hat{u}_t of p_t and u_t verify

$$z_t = \hat{p}_t + \hat{u}_t$$

The gain of the trend filter is presented in Figure 4, together with the gain function of the HP filter. Notice that the SEATS gain function is well above that of the HP filter, and is not a good approximation to the gain of the ideal filter that corresponds to our definition of the trend. According to this definition, we would obtain a decomposition of the form

$$z_t = \bar{z}_t + z_t^c$$

where \bar{z}_t is formed by all components of the series with period greater than eight years and z_t^c is the business cycle. Note that in this definition the components can and should be considered as the outcome of certain filters applied to the series. Specifically, let H_c be the ideal business cycle filter. Then, $z_t^c = H_c z_t$ and $\bar{z}_t = (1 - H_c)z_t$. It can be easily seen by inspecting Figure 4 that \bar{z}_t can be interpreted as a subcomponent of the trend estimator \hat{p}_t given by the model-based approach. We can argue that the model-based approach is not able to discriminate between \bar{z}_t and $\hat{p}_t - \bar{z}_t$ without further a priori information on the part of the user.

How can this a priori information be given to the procedure that should estimate such subcomponent? The answer is by means of a fixed filter H_n . This filter should be a trend filter (low-pass filter) such that, if applied to the component \hat{p}_t , would yield the smoother subcomponent $\bar{z}_t = H_n \hat{p}_t$. In the case in which the filter H_n is almost completely under the model-based filter, the filtered series $H_n \hat{p}_t$ will be practically identical to the series $H_n z_t$ obtained by filtering the original series. But if the series is, for example, white noise, then the model-based filter is zero and \bar{z}_t is zero as well. This procedure guarantees in an almost automatic manner that the fixed filter is correctly applied, since it takes into consideration the characteristics of the series. For the application, in the next section, we have first checked that it

is always the case that H_n is well under the model-based filter, so that we will estimate \bar{z}_t as $\bar{z}_t = H_n z_t$.

We emphasize that any kind of low-pass filters H_n can be used in the second step of the proposed procedure. For the empirical application, the following class H_n , which can be implemented in the program SEATS, will be used. Consider the following *ARIMA* model

$$\nabla \nabla_d z_t = (1 + \theta L) a_t, \quad (12)$$

where $\nabla_d z_t = (1 - L^d) z_t$. The canonical model-based decomposition that corresponds to this model is $z_t = \bar{z}_t - n_t$, where \bar{z}_t is the trend component and n_t is another component, orthogonal to \bar{z}_t . The trend component follows the model $\nabla^2 \bar{z}_t = (1 + L)(1 + \alpha L) b_t$ and the Wiener-Kolmogorov filter, denoted H_n , to obtain the minimum mean squared estimator $\hat{\bar{z}}_t$ is

$$\hat{\bar{z}}_t = H_n z_t = \frac{\sigma_b^2 (1 + L) \alpha(L) S(L) (1 + L^{-1}) \alpha(L^{-1}) S(L^{-1})}{\sigma_a^2 (1 + \theta L) (1 + \theta L^{-1})} z_t,$$

where $\sigma_a^2 = \text{Var}(a_t)$, $\sigma_b^2 = \text{Var}(b_t)$, $\alpha(L) = 1 + \alpha L$ and $S(L) = 1 + L + \dots + L^{d-1}$. The filter H_n is doubly infinite symmetric. It can be shown that

$$\frac{\sigma_b^2}{\sigma_a^2} = \frac{(1 + \theta)^2}{4(1 + \alpha)^2 d^2}, \quad k = \frac{2\theta}{(1 + \theta)^2} - \frac{d^2 + 2}{6}$$

and

$$\alpha = \frac{1 - k - (1 - 2k)^{1/2}}{k}$$

For each pair of parameters (θ, d) we obtain a trend filter H_n . For each pair (θ, d) , the gain function of H_n has value 1 at the zero frequency and decreases monotonically until it reaches the value zero at frequency $2\pi/d$. After that, it may present some small ripples between the frequencies $2k\pi/d$, $k = 1, 2, \dots, [d - 1]$.

In order to use one of these filters H_n in SEATS with yearly data, as shown in Figure 5, appropriate values for H_n are $d = 5$ and $\theta = -.6$. The resulting filter has less *leakage* and less *compression* than the HP filter for $\lambda = 10$.

3. An application to the GDP of OECD countries

3.1 Implementing the proposed model-based approach

In order to illustrate the use of the proposed model-based approach, we have conducted the following experiment. Instead of quarterly data, which are more limitedly available⁷, we used annual data for the OECD countries corresponding to the period 1960-1994, taken from OECD *National Accounts*. Since, as mentioned earlier, the performance of the filters at both ends of the series can be very poor, the sample was expanded first until 1995 using the rates of growth of GDP contained in the *Economic Outlook*. Then, using the automatic model identification facility of the program TRAMO, models were obtained and estimated for all countries in the sample and these models were used to forecast the series of GDP until the year 2005. The procedure used by the program also handles outliers automatically. The types of outliers considered are level shift, additive outliers and transitory changes. In all cases, we checked for model adequacy and we found that the fit was acceptable. Instead of using these same models for backcasting, we used the rates of growth of GDP implied in Summers and Heston's PWT 5.6 to backcast the series until 1950. Therefore, we used a total of 56 observations for each country to filter the series, although we were only interested in the filtered series for the period 1955-1995.

The models obtained by TRAMO were passed on to SEATS for signal extraction. For all countries, we checked that the trend filter obtained by SEATS using the canonical model-based decomposition was above the fixed filter H_n of the previous section corresponding to the values $d = 5$ and $\theta = -0.6$. Since this was the case for all countries, we decided to apply the fixed filter H_n directly to the series z_t , instead of applying it to the trend-cycle estimator \hat{p}_t obtained by SEATS. We decided not to correct for the outlier effects to facilitate the comparisons of the results with those obtained applying the HP

⁷ *Quarterly National Accounts* does not provide data for all OECD countries, and for most of them, quarterly GDP is only available from the beginning of the seventies.

filter.

3.2 Comparison between the HP filter and the proposed procedure

We also applied the HP filter, with a value of $\lambda = 10$, to the same GDP series of the previous section, using the same sample period.

In Table I, we present the standard deviation (σ) of the series filtered with both the HP ($\lambda = 10$) and SEATS ($d = 5, \theta = -.6$) filters as a measure of volatility. As can be seen, the estimates obtained with SEATS have a smaller standard deviation than those obtained with the HP filter in all cases. The values of σ given by this last filter are in line with the results obtained for quarterly data, using a value of $\lambda = 1600$, by, for example, Backus, Kehoe and Kydland (1993), from 1970 to mid nineties and for a smaller sample of countries. However, if the value $\lambda = 400$ is used for the annual US GDP, as proposed Hodrick and Prescott, a value of σ equal to 2.46 is obtained, which is higher than the typical value reported and confirms some of the problems pointed out in section 2. The last column of Table 1 shows the R^2 obtained by regressing the cyclical component given by the HP filter upon the one given by SEATS. As can be seen, the correlation between the two estimates of the business cycle is always very high, with Canada, Japan, Sweden and the United Kingdom being the countries with lower correlation.

Table 1
Cyclical Components of GDP: Some Statistics

Country	HP	Two-step	R^2
	σ	σ	
Australia	1.51	1.40	0.949
Austria	1.37	1.25	0.945
Belgium	1.34	1.06	0.944
Canada	1.67	1.21	0.896
Denmark	1.50	1.19	0.933
Finland	2.58	2.03	0.909
France	1.07	0.84	0.932
Germany	1.62	1.34	0.941
Greece	1.83	1.53	0.962
Iceland	3.21	2.41	0.927
Ireland	1.71	1.43	0.946
Italy	1.56	1.21	0.931
Japan	1.59	1.08	0.898
Luxembourg	2.01	1.76	0.947
Netherlands	1.63	1.43	0.953
New Zealand	2.41	1.98	0.901
Norway	1.21	0.92	0.905
Portugal	2.07	1.52	0.920
Spain	1.92	1.55	0.912
Sweden	1.26	1.02	0.896
Switzerland	2.06	1.57	0.926
Turkey	3.00	2.77	0.958
UK	1.49	1.19	0.900
USA	1.59	1.35	0.936

Notes: σ refers to the standard deviation of GDP cyclical components. The third column shows the R^2 of the regression of HP estimates upon SEATS ones for each country.

3.3 International comovements

To see the effect of the previous filters on the international comovements for the sample of countries considered, in Table 2 we present the correlation of each cyclical component for each country at time t with the cyclical component of the United States at times $t + k$ for $k = -1, 0, 1$. The results are similar for both filters and are in line with the evidence reported by Backus, Kehoe and Kydland (1993), or Baxter and Crucini (1993), for quarterly data, although the correlations we find are significantly lower for

Germany and Japan, and higher for the United Kingdom, probably because the data used by these authors are from 1970 onwards. The cross correlations obtained with the SEATS filter are higher than those obtained with the HP filter for 16 countries, which amounts to 2/3 of the sample. Comparing both results, differences seem to be important in the case of Denmark, Germany, Italy, Japan, New Zealand, Spain and Switzerland.

Table 2
International Comovements 1955-95
Correlation in t with cyclical componet of US GDP in $t+k$

Country	HP filter			Two-step Procedure		
	$k = -1$	$k = 0$	$k = 1$	$k = -1$	$k = 0$	$k = 1$
Australia	-0.04	0.46	0.31	-0.14	0.39	0.26
Austria	0.26	0.14	-0.13	0.34	0.26	-0.04
Belgium	0.28	0.20	-0.23	0.32	0.26	-0.26
Canada	0.46	0.74	0.13	0.35	0.68	-0.08
Denmark	0.01	0.46	0.23	0.04	0.54	0.14
Finland	0.36	0.23	-0.20	0.43	0.29	-0.22
France	0.43	0.23	-0.23	0.39	0.28	-0.21
Germany	0.15	0.34	-0.09	0.18	0.42	-0.08
Greece	0.15	0.28	0.19	0.03	0.28	0.25
Iceland	0.22	0.31	0.26	0.04	0.17	0.28
Ireland	0.23	0.30	-0.09	0.32	0.38	-0.10
Italy	0.47	0.23	-0.34	0.55	0.28	-0.42
Japan	0.18	0.10	-0.25	0.30	0.28	-0.19
Luxembourg	0.42	0.29	-0.25	0.47	0.37	-0.29
Netherlands	0.38	0.43	-0.10	0.34	0.41	-0.18
New Zealand	0.06	0.15	-0.14	0.20	0.23	-0.25
Norway	0.28	0.44	0.31	0.30	0.31	0.03
Portugal	0.32	0.24	-0.19	0.28	0.37	-0.07
Spain	0.37	0.10	-0.07	0.43	0.04	-0.17
Sweden	0.30	0.09	-0.19	0.28	-0.02	-0.37
Switzerland	0.38	0.27	-0.30	0.49	0.41	-0.30
Turkey	-0.24	-0.11	0.14	-0.33	-0.22	0.07
UK	0.40	0.71	0.29	0.29	0.73	0.23

3.4 Outliers

As mentioned earlier, the preceding comparison has been done without taking into consideration the effects of the possible outliers in the data, since *ad-*

hoc filtering does not allow for any treatment of outliers. However, it is possible that for some countries there exist some priors regarding possible transitory or permanent changes in their GDP, but *ad-hoc* filtering does not control for this possibility.

To show how the effects of outliers can be important, we present an example based in one well known particular experience. In 1959 Spain became a member of the IMF, which provided financial aid. This was done after Spain had taken a set of measures aimed at achieving fiscal and external equilibrium and at decreasing inflation. These measures had a positive long-run effects upon the GDP, but, in the short-run, the reduction in aggregate demand caused an important economic recession, which is now commonly interpreted as having had only transitory effects.

The GDP trend for the Spanish economy, for the period 1955-1965, estimated with the HP filter, can be seen in Figure 6. As already mentioned, the HP filter has a very long revision period. This is the reason why the trend level shows a lower rate of growth before 1960 and a higher one after this year. The series filtered with the SEATS filter, applied to the series corrected for the effect of the outlier, can also be seen in Figure 6. The difference between both estimates reaches 5 per cent of the GDP in 1960. Note that the effect of the outlier can be observed in the series filtered with the HP filter before the effect took place and that, because the outlier is transitory, it is assigned to the cyclical component estimated with the SEATS filter.

4. Conclusions

This paper has analyzed the properties of three well-known filtering methods to decompose economic variables into trend and cyclical components, following the engineering tradition of Butterworth filtering. It also compares the outcome of these methods with the ideal filter which corresponds to the standard and well accepted definition of the business cycle.

The HP filter has been shown to be a symmetrized Butterworth filter. This fact is important because it can be used to choose the smoothing parameter in terms of the frequency where the gain function is equal to $1/2$. The value of $\lambda = 1600$ proposed by Hodrick and Prescott (1997) for quarterly data is then justified. However, it suggests a value of $\lambda = 10$ for annual data, rather smaller than those frequently found in the empirical literature. Also, the use of forecasts and backcasts is recommended to improve the performance of the HP filter at both ends of the series.

In general, the HP filter performs reasonably well when applied to nonstationary seasonally adjusted quarterly data, corrected for outlier effects. This indicates that when all or any one of these characteristics do not apply, the use of the HP filter can be problematic. With annual data, the HP filter is a poorer approximation to the ideal filter, even using the smaller smoothing parameter $\lambda = 10$.

Compared with the HP filter, the BK procedure gives a better approximation to the ideal filter, but at the cost of more predictions at both ends of the sample. An additional advantage of the BK filter is that it can be applied to the original data, without using the controversial seasonally adjustment methods.

Ad-hoc filtering methods, like the HP and BK filters, face some important limitations. Firstly, there is no explicit definition for the trend and cyclical components. Secondly, it is impossible to make statistical inference, which is relevant in applications of *ad-hoc* filtering. Furthermore, these methods do

not consider specific treatment of outlier effects and, finally, *ad-hoc* filtering is an inappropriate method when, for example, applied to a stationary series.

A new method to obtain the BN decomposition has been proposed, which considerably simplifies the original procedure. Also, a recursive algorithm has been presented to obtain the BN estimates. With respect to the performance of the BN decomposition, it seems that the results do not improve those of the alternative methods.

To circumvent some of the problems associated with *ad-hoc* filtering, this paper proposes a new two-step method to obtain the trend and cyclical components. In the first step, a model-based procedure is used to obtain the trend-cycle signal. Then, in the second step, a fixed filter is applied to decompose this signal into the trend and cyclical components.

This procedure has several advantages. First, the trend-cycle component obtained in the first step depends on the characteristics of the series. Second, the series is corrected for outlier and other deterministic effects. Third, the fixed filter for annual data used in the second step approximates the gain of the ideal filter better than the HP filter and, finally, it requires fewer parameters than the BK filter. These improvements can be of importance when the statistics associated with Real Business Cycle Literature are computed to evaluate their performance or when the HP filter or other *ad-hoc* filtering methods are used to estimate the trend and cyclical components of economic time series.

To illustrate this point, the performance of the HP filter has been compared with that of the proposed procedure in an empirical application to estimate the business cycle of the GDP of OECD economies. To make the comparison possible, outlier treatment was not used. The results show that the cyclical components obtained with the two-step procedure present a smaller standard deviation in all cases. In addition, if transitory outliers are taken into account, the results obtained with the two-step procedure are significantly better.

All these results emphasize the importance of a more careful application of filtering in empirical literature and suggest avoiding one-fit-for-all solutions.

5. References.

- Backus, D. K. and P. J. Kehoe (1992): "International Evidence on the Historical Properties of Business Cycles". *American Economic Review*, 82, 864-888.
- Backus, D. K.; P. J. Kehoe and F. E. Kydland (1993): "International Business Cycles: Theory vs. Evidence". Federal Reserve Bank of Minneapolis, Quarterly Review, Fall, 14-29.
- Baxter, M. and M. J. Crucini (1993): "Explaining Saving-Investment Correlations". *American Economic Review*, 83(3), 416-36.
- Baxter, M. and R. G. King (1995): "Measuring Business Cycles: Approximate Band-pass Filters for Economic Times Series". NBER WP 5022.
- Bell, W. (1984): "Signal Extraction for Nonstationary Series". *Annals of Statistics*, 12, 646-664.
- Beveridge, S. and C. R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'". *Journal of Monetary Economics* 7, 151-174.
- Burman, J. P. (1980): "Seasonal Adjustment by Signal Extraction". *Journal of the Royal Statistical Society A*, 143, 321-337.
- Cogley, T. and J. M. Nason (1995): "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series. Implications for Business Cycle Research". *Journal of Economic Dynamics and Control*, 19, 253-278.
- Correia, I., Neves, J.C. and Rebelo, S. (1995): "Business Cycles in a Small Open Economy". *European Economic Review*, 39, 1089-1113.
- Gersch, W. and G. Kitagawa (1988): "Smoothness Priors in Time Series", in J. C. Spall (ed.), *Bayesian Analysis of Time Series and Dynamic Models*. Marcel Dekker, Inc. New York.
- Gómez, V. and A. Maravall (1997). "Program TRAMO and SEATS. Instructions for the User". Working Paper SGAPE-97001. DGAPP. Ministry of Economy and Finance. Madrid.
- Hansen, G.D. and Wright, R. (1994): "The Labor Market in Real Business Cycle Theory", in P.J. Miller (ed.) *The Rational Expectations Revolution*. Massachusetts Institute of Technology.
- Harvey, A. C. and A. Jaeger (1993): "Detrending, Stylized Facts and the Business Cycle". *Journal of Applied Econometrics*, 8, 231-247.

- Hodrick, R. and E. C. Prescott (1997): "Post-war US Business Cycles: An Empirical Investigation". *Journal of Money, Credit and Banking*, 29(1), 1-16.
- King, R. G. and S. T. Rebelo (1993): "Low Frequency Filtering and Real Business Cycles". *Journal of Economic Dynamics and Control*, 17, 207-231.
- Kohn, R. and C. F. Ansley (1988): "Equivalence Between Bayesian Smoothness Priors and Optimal Smoothing for Function Estimations", in J. C. Spall (ed.), *Bayesian Analysis of Time Series and Dynamic Models*. Marcel Dekker, Inc. New York.
- Kydland, F. and Prescott, E.C. (1982): "Time to Build and Aggregate Fluctuations". *Econometrica*, 50, 1345-1370.
- Maravall, A. (1993): "Stochastic Linear Trends. Models and Estimators". *Journal of Econometrics* 56, 5-37.
- Prescott, E. C. (1986): "Theory Ahead of Business Cycle Measurement", in K. Brunner and A.H. Meltzer (ed.) *Real Business Cycles, Real Exchange Rates and Actual Policies*. Carnegie-Rochester Conference Series on Public Policy, vol. 25, pp. 11-24.
- Stadler, G. W. (1994): "Real Business Cycles". *Journal of Economic Literature*, 32(4), 1750-83.
- Summers, R. and A. Heston (1991): "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988". *Quarterly Journal of Economics*, 106(2), 327-368.
- Whittaker, E. T. (1923): "On a New Method of Graduation". *Proc. Edinburgh Math. Assoc.*, 78, 81-89.

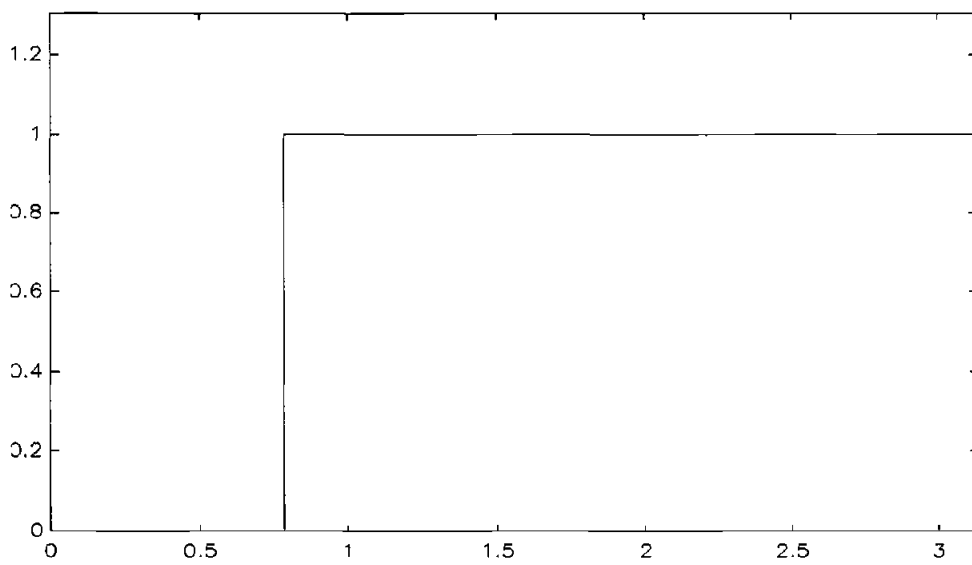


Figure 1: *Gain of the ideal filter (yearly data).*

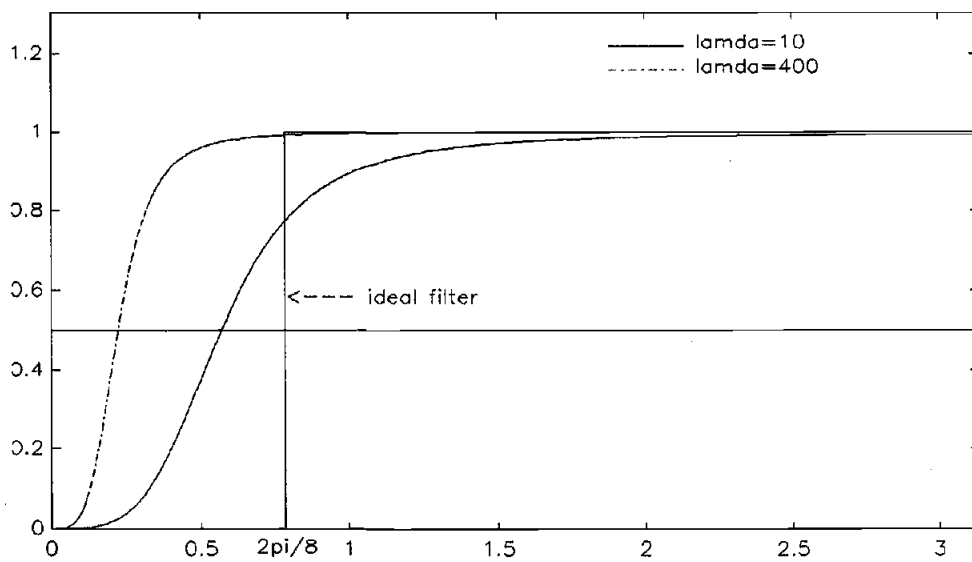


Figure 2: Gain of the HP filter with anual data

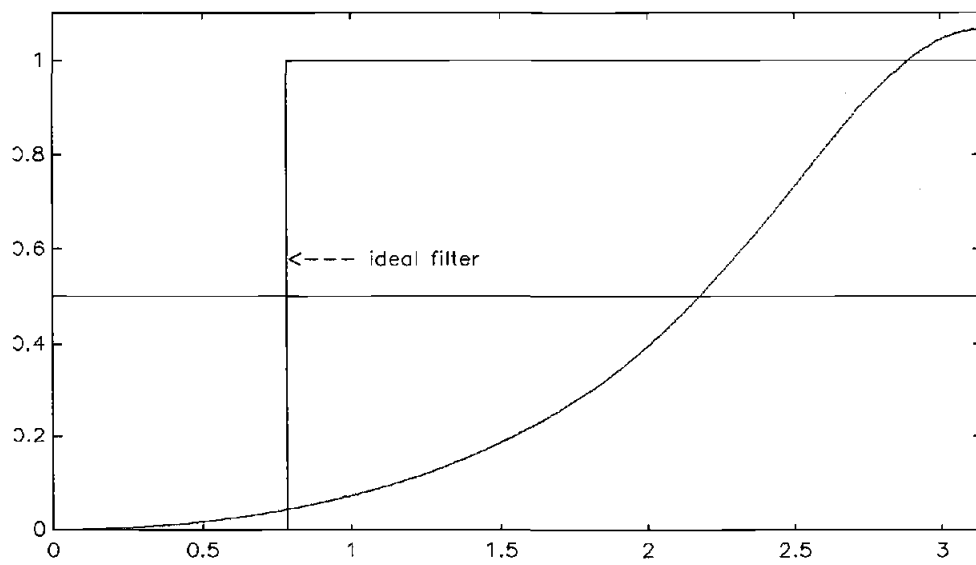


Figure 3: Gain of the Beveridge-Nelson filter

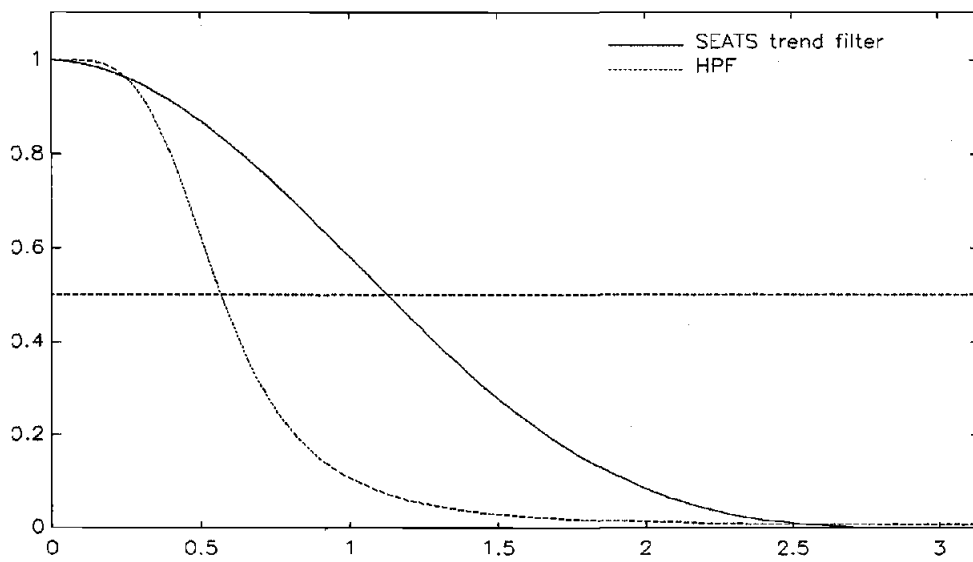


Figure 4: Gain of the trend of the HP and SEATS filters.

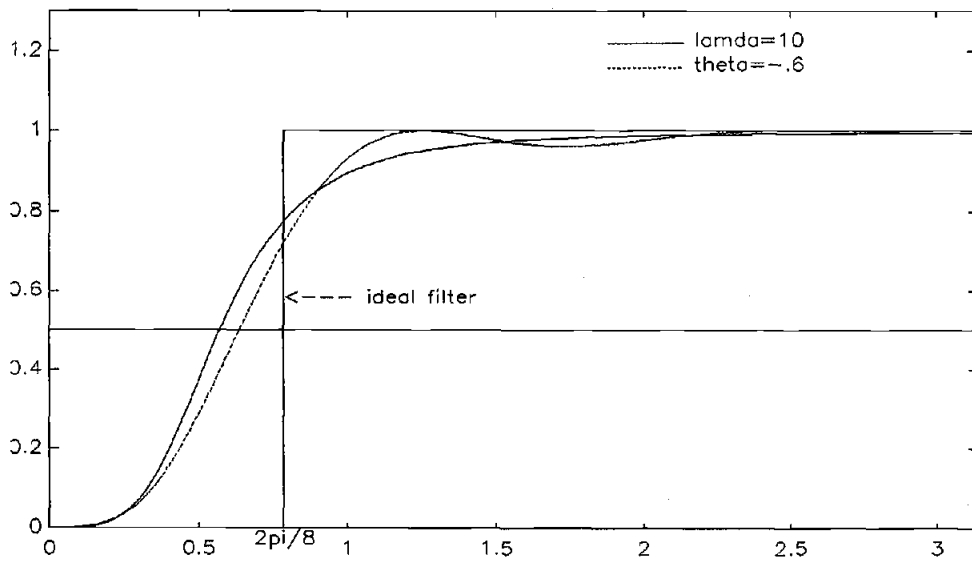


Figure 5: Gain of the HP ($\lambda = 10$) and the SEATS filters ($d = 5, \theta = -0.6$)

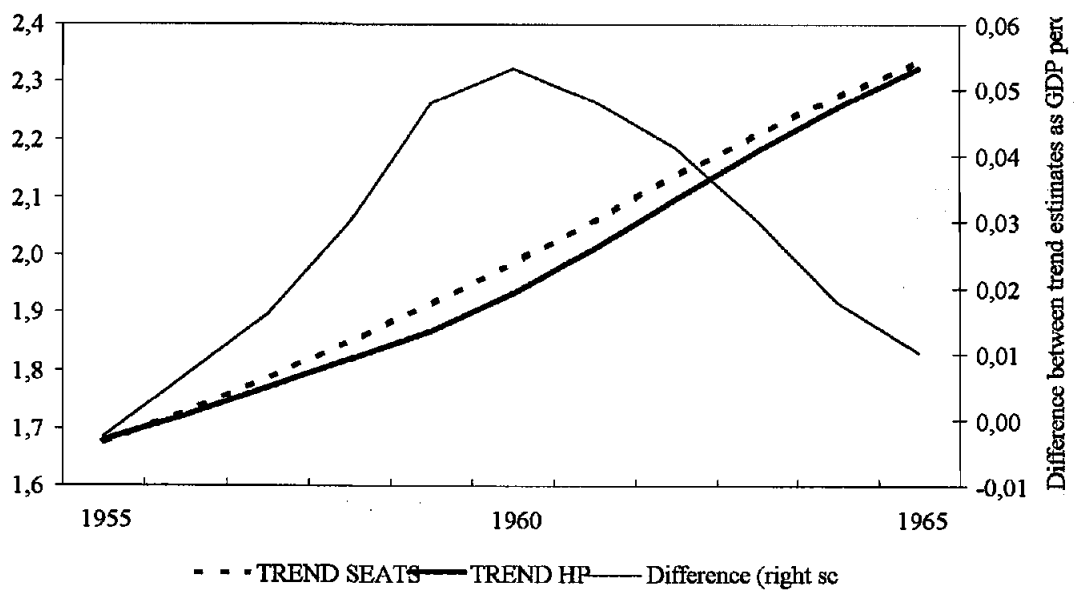


Figure 6: Trend estimates for the Spanish economy from 1955 to 1965.