



A Frequency Selective Filter for Short-Length Time Series

**N° 2004-05
May 2004**

Alessandra IACOBUCCI
OFCE

A Frequency Selective Filter for Short-Length Time Series*

Alessandra Iacobucci¹ and Alain Noullez²

¹ *OFCE, 69, quai d'Orsay, F-75340 Paris Cedex 07*

¹ *CNRS – IDEFI, 250 rue Albert Einstein, F-06560 Valbonne, France*

`alessandra.iacobucci@sciences-po.fr`

² *Observatoire de Nice, B.P. 4229, F-06304 Nice Cedex 4, France*

`anz@obs-nice.fr`

Abstract

An effective and easy-to-implement frequency filter is designed by convolving a Hamming window with the ideal rectangular filter response function. Three other filters (Hodrick-Prescott, Baxter-King, and Christiano-Fitzgerald) are critically reviewed. The behavior of the Hamming-windowed filter is compared to the others through their frequency responses and by applying them both to an artificial, known-structure series and the Euro zone GDP quarterly series. As for the Hodrick-Prescott filter, a bandpass version of it is used. The Hamming-windowed filter has almost no leakage and is thus much better than the others at eliminating high frequency components, while the response in the passband is significantly flatter. Moreover, its behavior at low frequencies ensures a better removal of undesired long-term components. Those improvements are particularly evident when working with short-length time series, which are common in Macroeconomics. The proposed filter is stationary, symmetric, uses all the information contained in the raw data and stationarizes series integrated up to order two. It thus proves to be a good candidate for extracting frequency-defined business cycle components.

Keywords: Spectral methods, frequency selective filters, Hodrick-Prescott, Baxter-King and Christiano-Fitzgerald bandpass filters, business cycles.

JEL codes: C10, C14, E32.

*A. Iacobucci is grateful to Andrew Harvey and Simon van Norden for a challenging discussion on the occasion of the EUROSTAT colloquium (Luxembourg, October 2003). She thanks Guillaume Chevillon, Matthieu Lemoine and Francesco Saraceno for their precious comments and suggestions.

1 Introduction

There are several ways of formalizing the separation of a signal into different periodic components. One of the most insightful remains the Fourier decomposition, which views the signal as a linear combination of purely periodic components, each having an amplitude constant in time and a well-defined frequency. Short-time Fourier analysis and wavelets (which make it possible to represent the frequency content of a series, while keeping the time description parameter) are also an alternative, especially in the case of nonstationary or intermittent signals. These techniques allow a detailed insight of the data structure. However, they are not easily implementable with short-length series and will not be broached in this paper.

The selective filtering operation over an infinite continuous signal is defined by specifying the range of individual frequencies that should be extracted and those that should be removed. In the case of finite-length samples, it is impossible to design such a filter, preserving all frequencies in a given range and completely removes those outside it (the so called *ideal filter*). Indeed, abrupt variations of the frequency response give rise, especially in the case of short-length series, to the *Gibbs phenomenon*, i.e. the appearance of spurious artificial fluctuations in the filtered signal. Therefore, it is important to design good approximations of the ideal filter, “good” referring to some optimization criteria, like the (weighted) difference between the desired and the effective response.

In this paper, a new filter is proposed, the HW filter, which provides a simple and efficient solution to the ideal filter approximation problem. This filter is obtained by smearing the ideal filter response with a lag window and it leads to a very good attenuation of the spectral power outside the passband, allowing the almost complete removal of undesired frequency components. The only drawback is a negligible widening of the transition between passband and stopband¹, which means that a negligible part of the frequency components lying near the edge of the chosen band may be present in the filtered series.

In order to show its qualities and improvements, we compare the HW filter to those most widely used in Macroeconomics for trend and cycle extraction, namely the filters by Hodrick and Prescott [9], Baxter and King [1] and Christiano and Fitzgerald [3]. We thoroughly review and discuss these filters: in particular, we visualize the time coefficients, the frequency response and the phase of the Christiano-Fitzgerald filter, by plotting them as surfaces in three-dimensional space.

Before proceeding, we would like to spend a word in favor of the nonparametric filtering technique, which, after a period of splendor, has lately come under attack (see, among others, [2, 12]). For instance, Murray [12] shows that the Baxter-King filter, and in general any bandpass filter, is unable to extract the business cycle from an unobserved components model with a $I(1)$ trend. The first response to this criticism is that it is obvious from the outset that no spectral filtering procedure alone can isolate the cyclical component in such a model. In fact, it is well-known that the spectrum of an $I(1)$ process goes like ν^{-2} at small frequencies. Thus the $I(1)$ trend in Murray’s model has non-zero frequency components spread over the whole spectrum. Their amplitude — which is likely to predominate especially at low and very low frequencies —, inevitably adds to that of the cycle. We would then need a tool to separate the different contributions *within* the individual frequency component, something a bandpass

¹For the sake of clarity, the terms *passband* and *stopband* are taken from signal analysis terminology and refer to “the chosen band” and “all the rest” of the spectrum, respectively.

filter is not meant for. Secondly, if we have no prior information on the process generating the data, the “non-structural” [2] method of spectral analysis is just as good as the model-based approach in extracting business cycles: it is just a matter of different business cycle definitions. Furthermore, we need not forget that it is almost always the case that more than one model can satisfactorily fit the same raw data. And even when one of them is clearly better than the others, it is often thanks to *ad hoc* hypotheses that make the whole machine work. To sum up, nonparametric spectral filtering, as all other methods in time series analysis, has limits which have to be known and thoroughly explored to ensure a proper utilization; and, as for any other method or model, we can not expect it to be infallible and universal.

The paper is organized as follows: in the following section the windowed filter is introduced, together with a sketch of the computing algorithm. The third section contains an extensive critical overview of the the Hodrick-Prescott (HP), the Baxter-King (BK) and the Christiano-Fitzgerald (CF) filter. In the fourth section these are compared to our windowed filter, both from a theoretical point of view by plotting their frequency responses and from an applied one, by applying them to an artificial series and the Euro zone GDP quarterly series. The fifth section concludes.

2 The Windowed Filter

Consider the filtering problem for a finite time series u_j of duration $T = N\Delta t$, where N is the number of data points and Δt the sampling periodicity. The simple truncation of the ideal filter time coefficients

$$\begin{aligned} h_j^{\text{ideal}} &= \frac{\sin(2\pi\nu_h j \Delta t) - \sin(2\pi\nu_l j \Delta t)}{\pi j}, & j = -\infty, \dots, \infty, \\ h_0^{\text{ideal}} &= \lim_{j \rightarrow 0} h_j^{\text{ideal}} = 2\Delta t(\nu_h - \nu_l), \end{aligned} \tag{1}$$

where $j = -\infty, \dots, \infty$ and ν_h and ν_l represent the high and low cutoff frequency respectively, adds up to multiplying (1) by an N -wide rectangular lag window. This yields

$$\begin{aligned} h_j &= \frac{\sin(2\pi k_h j / N) - \sin(2\pi k_l j / N)}{\pi j} & j = 1, \dots, N - 1, \\ h_0 &= \frac{2(k_h - k_l)}{N} \end{aligned} \tag{2}$$

where we have set $\nu_l = k_l / (N\Delta t)$ and $\nu_h = k_h / (N\Delta t)$. The truncation produces fluctuations of large amplitude and slow decay in the response function. This is caused by the discontinuities induced by the lag window, whose $\sin(\pi\nu T) / (\pi\nu T)$ -profile Fourier transform (Figure 1, right panel) disturbs the ideal frequency response. An adjustment of the rectangular window shape is thus required to obtain a response that goes faster to zero. For this purpose, the “adjusting” window should be chosen to go to zero continuously with its highest possible order derivatives, at both ends of the observation interval. The choice of the best window has been thoroughly discussed in signal processing literature (see e.g [10, 13]) and is not a univocally defined problem because different optimization criteria could be used, like the minimization of the total energy outside of the main lobe (the highest peak in Figure 1, right panel), or of the maximum

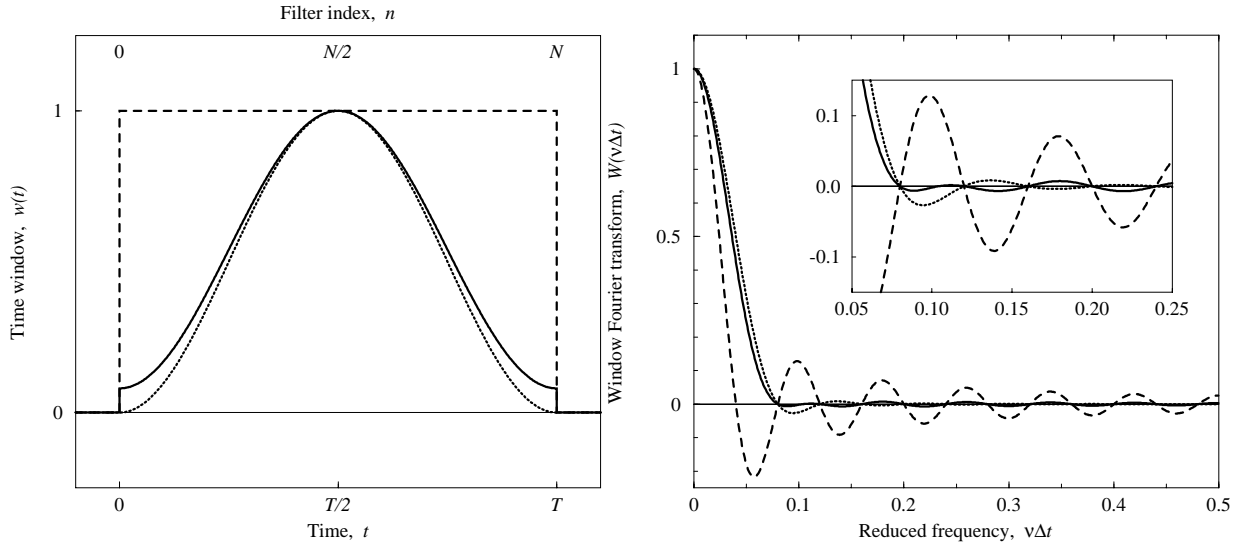


Figure 1: **Window Functions and Their Frequency Response.** *The rectangular (dashed line), Hanning (dotted line) and Hamming (full line) time windows (left panel) and their respective spectral windows (right panel). Note in the zoom (right panel, inset) the reduced side lobe amplitude and leakage of the Hanning and Hamming windows with respect to the rectangular one, the Hamming window performing better in the first side lobe.*

amplitude over all side lobes or of the amplitude of the first one. Examples of lag windows widely used in spectral analysis are, for instance, those by Bartlett, Parzen, von Hann (the so-called Hanning window) and Hamming. We considered these last two, since they were likely to meet our requirements. The Hanning lag window

$$w_j^{\text{Han}} = \frac{1}{2} - \frac{1}{2} \cos\left(\frac{2\pi j}{N}\right) \quad (3)$$

is continuous and vanishes at both zero and N along with its first derivative (see Figure 1, left panel). Its corresponding spectral window (i.e. its Fourier transform)

$$\begin{aligned} W^{\text{Han}}(\nu) &= \frac{1}{2} \frac{\sin(\pi\nu T)}{\pi\nu T} - \frac{1}{4} \left\{ \frac{\sin[\pi(\nu T + 1)]}{\pi(\nu T + 1)} + \frac{\sin[\pi(\nu T - 1)]}{\pi(\nu T - 1)} \right\} \\ &= \frac{1}{1 - (\nu T)^2} \frac{\sin(\pi\nu T)}{\pi\nu T} \end{aligned} \quad (4)$$

has a first side lobe amplitude of only $2.8 \cdot 10^{-2}$ and goes to zero like $(\nu T)^{-3}$ in the high-frequency zone, as shown in Figure 1. We recall that, because of discretization, the absolute value of ν is bounded by the so-called *Nyquist frequency*, $\nu_{\text{Nyq}} = (2\Delta t)^{-1}$.

The Hamming lag window

$$w_j^{\text{Ham}} = 0.54 - 0.46 \cos\left(\frac{2\pi j}{N}\right) \quad (5)$$

is obtained by a judicious combination of the Hanning and the rectangular lag window (Figure 1). The Hamming window is not continuous at the edges of the interval and its corre-

spending spectral window

$$\begin{aligned} W^{\text{Ham}}(\nu) &= 0.54 \frac{\sin(\pi\nu T)}{\pi\nu T} - 0.23 \left\{ \frac{\sin[\pi(\nu T + 1)]}{\pi(\nu T + 1)} + \frac{\sin[\pi(\nu T - 1)]}{\pi(\nu T - 1)} \right\} \\ &= \left[0.08 - \frac{0.92}{1 - (\nu T)^2} \right] \frac{\sin(\pi\nu T)}{\pi\nu T} \end{aligned} \quad (6)$$

decreases like $(\nu T)^{-1}$ for large ν , but with a much smaller amplitude than the rectangular window (see Figure 1). Thus, the Hanning spectral window performs better than the Hamming spectral window in the upper part of the spectrum (high ν), and this is useful in the case of long time series with a good frequency resolution, like those typical, for example, in finance. However, since it solves the optimization problem which minimizes the amplitude of the first side lobe, the Hamming spectral window has a first side lobe amplitude of only $8.9 \cdot 10^{-3}$, thus quite reduced with respect to that of the Hanning spectral window. Hence, it is the appropriate window to use when frequencies close to the edges of the passband have to be eliminated. This window is thus more appropriate to macroeconomic time series, especially for the extraction of the cycle components.

Both Hanning and Hamming spectral window are real and even, so that the symmetries of the ideal infinite filter are preserved, that is, if the latter is real and even, it remains so through windowing. They have a main lobe which is twice as wide as the one for the rectangular window. This implies that the filters obtained by windowing will have a transition band approximately twice as wide as that of the ideal filter — the price to pay for the smoothing.

From the expression of the filter frequency response

$$H(\nu) = \Delta t \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} h_n e^{i2\pi\nu j \Delta t}, \quad (7)$$

we see that the sum of the windowed filter coefficients h_j are set by the frequency response at the origin

$$H(0) = \Delta t \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} h_n. \quad (8)$$

This allows the removal of the signal mean

$$H(\nu)|_{\nu=0} = 0, \quad (9)$$

in the case of lowpass and bandpass filters. If we add the condition

$$\left. \frac{dH(\nu)}{d\nu} \right|_{\nu=0} = 0, \quad (10)$$

or its time domain equivalent

$$\sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} n h_n = 0, \quad (11)$$

the elimination of two unit roots is ensured. Both Hamming and Hanning windowed filters satisfy these conditions, thus they stationarize $I(2)$ series.

If the series has a deterministic trend, for instance a polynomial function of time, the subtraction of ordinary least-squares fit must be performed before filtering. Indeed, in this case the operation of detrending and filtering do not commute². Subtracting the OLS regression line can always be done to reduce the edges discontinuity that would appear if there were non-harmonic frequencies in the signal. The effect of this discontinuity is however reduced to negligible amplitudes by the use of windowing.

2.1 The HW Filter Algorithm

Let us come to the introduction of the proposed filtering procedure, described by the following algorithm³:

- (i) subtract, if needed, the least-square line to remove the artificial discontinuity introduced at the edges of the series by the Fourier transform;
- (ii) compute the discrete Fourier transform of the signal u_j

$$U_k = \frac{1}{N} \sum_{j=0}^{N-1} u_j e^{-i2\pi jk/N}, \quad k = 0, \dots, \lfloor N/2 \rfloor ;$$

- (iii) apply the Hanning- or Hamming-windowed filter ($W_k * H_k$) to U_k

$$\begin{aligned} V_k &= (W_k * H_k) U_k = \sum_{k'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} W_{k'} H_{k-k'} U_k \\ &= (0.23 H_{k-1} + 0.54 H_k + 0.23 H_{k+1}) U_k, \quad k = 0, \dots, \lfloor N/2 \rfloor, \end{aligned}$$

where H_k is defined by the frequency range as

$$H_k = H_k^{\text{ideal}} \equiv \begin{cases} 1 & \text{if } \nu_l N \Delta t \leq |k| \leq \nu_h N \Delta t \\ 0 & \text{otherwise} \end{cases},$$

and W_k is given by substituting $\nu = k/T$ in Equation (4) or (6);

- (iv) compute the inverse transform

$$v_j = \left[V_0 + \sum_{k=1}^{\lfloor N/2 \rfloor} \left(V_k e^{i2\pi jk/N} + V_k^* e^{-i2\pi jk/N} \right) \right], \quad j = 0, \dots, N-1.$$

Windowing the filter response in the frequency domain by convolution of the ideal response with the spectral window, as in (iii), is computationally more profitable to perform than the time domain multiplication. Indeed, both Hamming and Hanning spectral windows have only three non-zero components, at frequencies 0 and $\pm(N\Delta t)^{-1}$.

From now on we will deal only with the Hamming-windowed (HW) filter, since is more suitable for business cycle extraction. By the way, the differences between the two filters in the comparison with the others would be almost imperceptible. Figure 2 shows the frequency

²As a matter of fact, the decomposition of a polynomial in the Fourier basis is not unique since the former contains all possible frequency components. Consequently, a polynomial and a periodic term are not orthogonal.

³A `gawk` (<http://www.gnu.org/software/gawk/gawk.html>) program of the filter is available on request. A matlab version will be soon available.

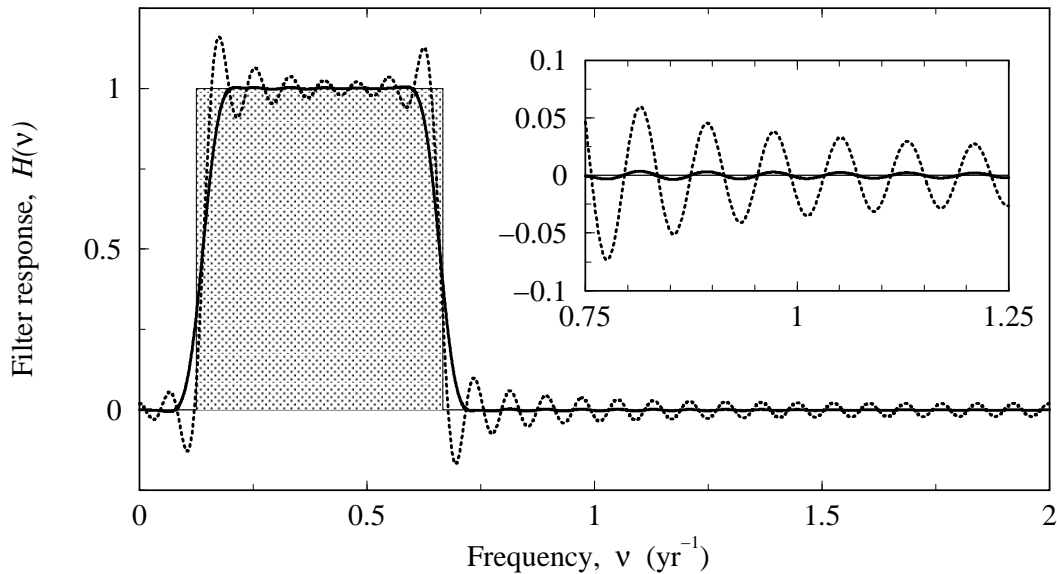


Figure 2: **Windowed Filter Frequency Response.** *By convolving the signal with a Hamming window, a filter frequency response with strongly reduced leakage and compression can be obtained. Note the negligible amplitude of the extra lobes of the frequency response as shown in the zoom (inset).*

response of bandpass filters obtained by direct truncation and by the Hamming-windowing procedure. Note the reduced leakage of the filter obtained by windowing. As for the wider transition band, however, in many applications it is more important to remove the most part of undesired frequencies than to have a very sharp discrimination between frequencies at the edges of the passband.

To sum up, this procedure ensures the best possible behavior in the upper part of the spectrum, the complete removal of the signal mean and a flat response inside the passband.

3 Comparison with Other Filters

In this section, we consider some filters widely used in the literature for trend and cycle extraction, namely the HP filter [9], the BK filter [1] and CF filter [3]; after briefly describing them, we will make a comparison with our filter. We compute the bandpass version of the HP filter in order to make a fairer comparison.

3.1 The Hodrick-Prescott Filter

The Hodrick-Prescott cyclical component v_j^{HP} is defined as the difference between the original signal u_j and a smooth growth component g_j . The latter is the solution of the optimization problem

$$\min_{\{g_j\}} \sum_{j=0}^{N-1} \left[(u_j - g_j)^2 + \lambda (g_{j+1} - 2g_j + g_{j-1})^2 \right], \quad (12)$$

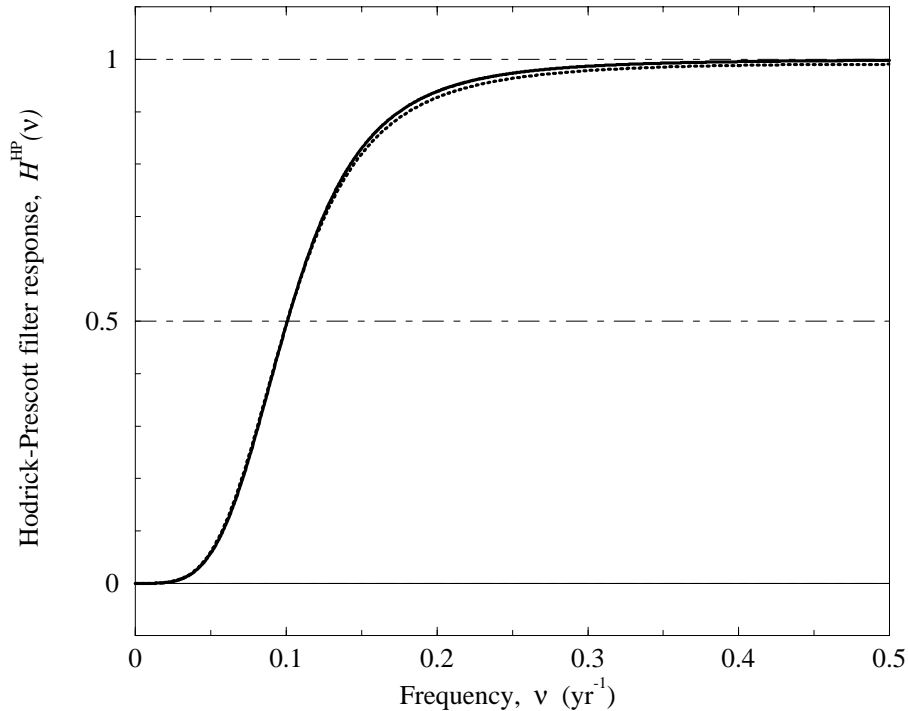


Figure 3: **Hodrick-Prescott Filter Frequency Response.** Frequency response of the Hodrick-Prescott filter using $\lambda = 1600$ applied to quarterly data (full line) (in the frequency range $[0, 0.5 \text{ yr}^{-1}]$), and of a Hodrick-Prescott filter using $\lambda = 6.68$ applied to annual data (dotted line).

which minimizes the sum of norm of the cyclical component $\|v_j^{\text{HP}}\| = \|u_j - g_j\|$ and the weighted norm of the rate of the growth component $\|(1 - L)(1 - L^{-1})g_j\|$, where L is the lag operator $Lx_j \equiv x_{j-1}$. The smoothing parameter λ penalizes variations in the growth rate with respect to the differences between filtered and unfiltered series and is usually set to 1600 for quarterly data. For large values of λ , the growth component g_j tends to the OLS line calculated on the data.

The solution of (12) can be found explicitly in the frequency domain [11] and leads to the following expression for the frequency response function

$$\begin{aligned}
 H^{\text{HP}}(\nu) &= \frac{V^{\text{HP}}(\nu)}{U(\nu)} = \frac{U(\nu) - G(\nu)}{U(\nu)} \\
 &= \frac{4\lambda(1 - \cos(2\pi\nu\Delta t))^2}{1 + 4\lambda(1 - \cos(2\pi\nu\Delta t))^2} \\
 &= \frac{16\lambda \sin^4(\pi\nu\Delta t)}{1 + 16\lambda \sin^4(\pi\nu\Delta t)}, \tag{13}
 \end{aligned}$$

where $G(\nu)$ is the Fourier transform of the series growth component g_j . From this expression and Figure 3, it can be seen that this is in fact a highpass filter, the frequency response rising monotonically from zero at $\nu = 0$ to nearly one at the Nyquist frequency. The transition is rather smooth and occurs at a cutoff frequency — defined as the frequency for which the

response is equal to 0.5 — given by

$$\nu_c = (\pi\Delta t)^{-1} \arcsin\left(\frac{\lambda^{-1/4}}{2}\right), \quad (14)$$

i.e. $\nu_c = 0.0252(\Delta t)^{-1}$ when $\lambda = 1600$. Hence the HP filter, in the configuration suggested by the authors for quarterly data, selects periodicities shorter than 10 yr approximately. The HP filter is real and symmetric, so that its frequency response is also real and symmetric and introduces no phase shift, but the filter response has the disadvantage of a wide transition band (see Figure 3). The frequency response goes like $\lambda(2\pi\nu\Delta t)^4$ at small frequencies, hence it behaves as a fourth-difference filter and can stationarize $I(4)$ processes.

A recurring issue when using the HP filter is the value of the parameter λ to use when dealing with annual or monthly data. This has been studied, among others, by Ravn and Uhlig [14], who found

$$\lambda_s = s^4 \lambda_q, \quad (15)$$

where λ_q is the value of the parameter for quarterly data, i. e. the usual 1600, s is the alternative sampling frequency (equal to 1/4 for annual data and 3 for monthly data) and λ_s the corresponding parameter value. As for the dependence of λ on the cutoff frequency, by noticing that the HP filter belongs to the wider class of Butterworth filters, Gomez [5] found the expression

$$\lambda = [2 \sin(x_c/2)]^{-4}, \quad (16)$$

where x_c stands for the reduced angular frequency. A more general and comprehensive work has been done by Harvey and Trimbur [8]. They analyze the dependence of λ on cutoff frequency and sampling frequency in a model-based approach to filters. In particular they are interested in how the variation of λ can change the structure of an unobserved component model, by modifying, for example, the correlation between components.

To find the explicit dependence of λ on both the sampling frequency and the cutoff frequency, we can directly solve the frequency response (13) for λ . Through (14), we obtain

$$\lambda = [2 \sin(\pi\nu_c\Delta t)]^{-4}. \quad (17)$$

Consequently, for a value of ν such that $\lambda = 1600$ for $\Delta t = 4^{-1}$ yr, we obtain that $\lambda = 6.68$ when $\Delta t = 1$ yr (see Figure 3) and $\lambda = 129660$ when $\Delta t = 12^{-1}$ yr for such that. Equations (16) and (17) are exactly the same, considering that $x_c = 2\pi\nu_c\Delta t$. It is easy to see that, in the limit of small values of the reduced frequency $\nu_c\Delta t$ (i.e. large values of λ), the formula (15) is equivalent to (17) and of course (16), and both equivalent to

$$\lambda \approx (2\pi\nu_c\Delta t)^{-4} + O\left((\nu_c\Delta t)^{-2}\right). \quad (18)$$

Therefore λ varies as the inverse fourth power of ν_c or Δt , as could be guessed by inspection of (12), where λ multiplies the square of the second difference of g_j .

By means of the formula (17), or (18), which allows the tuning of the cutoff frequency, a Hodrick-Prescott *bandpass* filter can be obtained as the difference between two highpass filters with different appropriate λ values, as shown in [5] and [7] for Butterworth filters. Figure 4 shows the [2, 8] yr bandpass HP filter obtained for samples of $N = 33$ (top panel)

and $N = 128$ (bottom panel) quarterly data points. Since the length⁴ of this filter is equal to the total number of data points, the behavior of the filter improves as the number of data available grows. However, as expected, the bandpass HP filter cumulates the compression of two standard HP filters and is thus a quite poor approximation of the ideal bandpass filter.

3.2 The Baxter-King Filter

The method proposed by Baxter and King [1] relies on the use of a symmetric finite odd-order $M = 2K + 1$ moving average so that

$$\begin{aligned} v_j &= \sum_{n=-K}^K h_n u_{j-n} \\ &= h_0 u_j + \sum_{n=1}^K h_n (u_{j-n} + u_{j+n}) . \end{aligned} \quad (19)$$

The set of M coefficients $\{h_n^{\text{BK}}\}$ is obtained by truncating the ideal filter coefficients at M under the constraint of the correct amplitude at $\nu = 0$, i.e. $H(0) = 0$ for bandpass and highpass filters and $H(0) = 1$ for lowpass filters (Equation (8)). The BK filter coefficients have thus to solve the following optimization problem

$$\begin{aligned} \min_{\{h_n^{\text{BK}}\}_{n=-K, \dots, K}} & \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} d\nu \left| \sum_{n=-K}^K (h_n^{\text{BK}} - h_n^{\text{ideal}}) e^{-i2\pi n\nu\Delta t} \right|^2 \\ \text{s.t.} & \sum_{n=-K}^K h_n^{\text{BK}} = \frac{H(0)}{\Delta t} . \end{aligned} \quad (20)$$

The solution of the constrained problem simply involves shifting all ideal coefficients by the same constant quantity

$$h_j^{\text{BK}} = h_j^{\text{ideal}} + \frac{H(0) - \Delta t \sum_{n=-K}^K h_n^{\text{ideal}}}{M\Delta t} . \quad (21)$$

The frequency response of the BK filter with $K = 16$ selecting the band $[2, 8]$ yr is reported in Figure 4, together with the responses of the other filters examined in this paper. Since the length of the BK filter is always $M = 2K + 1$ and does not depend on N , we first plot the responses of all the filters we are considering for $N = 33$ points of quarterly data. In this way we can compare filters of equal length. Then we plot the same curves for 128 quarterly data points, to show how the N -length filters improve with growing values of N , while the BK filter remains unchanged (as does the fixed-length symmetric version of the CF filter, see next section).

Beside being optimal for the constrained problem (20), the BK filter has many desirable properties. First, since it is real and symmetric, it does not introduce any phase shift and leaves the extracted components unaffected except for their amplitude. Second, being of constant finite length and time-invariant, the filter is stationary. Third, the filter is symmetric and

⁴The *length* of a filter is the number of points involved in the calculation of one point of the filtered series.

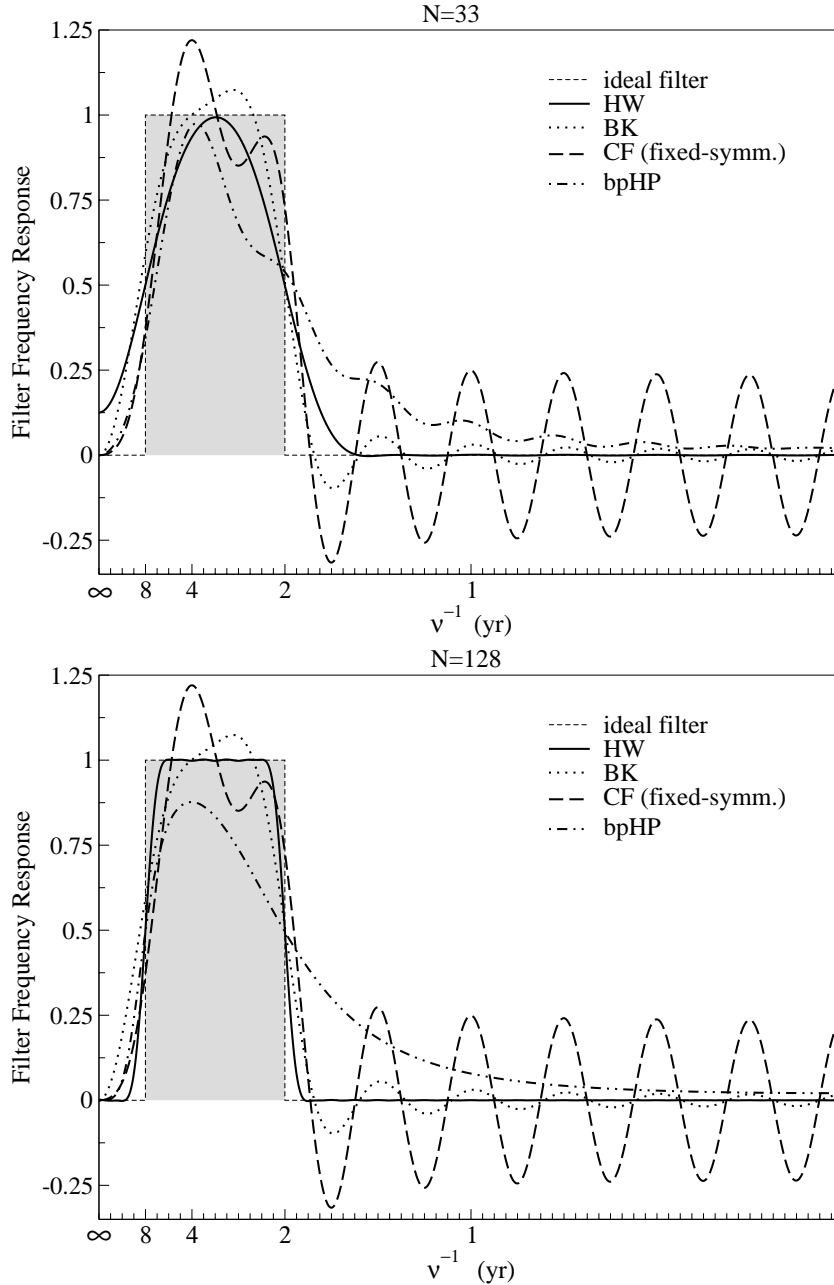


Figure 4: **Comparison of Different Filters.** Frequency responses of the four different approximate bandpass filters selecting periods between 2 yr (8 quarters) and 8 yr (32 quarters) using 33 (top panel) and 128 (bottom panel) points of quarterly data. These filters are: the proposed HW filter (full line); the BK filter with $K = 16$, as suggested in [1] (dotted line); the CF random walk filter in the symmetric, fixed-length ($K=16$) version (dashed line, see text for details); the bandpass HP filter obtained as the difference between the two standard (highpass) HP filter with $\lambda_{low} = 677.1298$ (8 years) and $\lambda_{high} = 2.9142$ (2 years).

satisfies (8) thus correctly eliminates the signal mean. Moreover the bandpass and highpass filter response behave at least like ν^2 for small ν which allows to remove up to two unit roots (see (9) and (10)). Fourth, the filter is insensitive to deterministic linear trends, provided

that $M < N$, i.e. that it is not used near the edges of the series.

On the other hand, filtering in time domain using moving averages, involves the loss of $2K$ data values. On the other hand, if too small a value of K is chosen, the filter resolution $[(2K + 1)\Delta t]^{-1}$ would worsen. In [1], a value of $K = 12$ for the passband $[2, 8]$ yr is found to be equivalent to higher values, such as 16 or 20. Thus the authors suggest to put $K \geq 12$ irrespectively of N , of the sampling frequency Δt and of the band to be extracted. This may cause significant compression and high leakage in the obtained filter response. Such drawbacks are of course consequences of the truncation, but they are undoubtedly amplified by the constraint imposition in (20), which, by adding a constant value to the ideal filter coefficients (see Equation (21)), causes an extra discontinuity at the endpoints of the filter, worsening the leakage at high frequencies.

In our opinion, the correct procedure is to take into account the filter resolution, once we have fixed the cutoff frequencies. In fact, the value of M must be such that $\nu_h - \nu_l > (M\Delta t)^{-1}$, otherwise the filter is unable to select the band with enough accuracy. In particular, filters with $\nu_l > (M\Delta t)^{-1}$ — i.e. filters whose length M is smaller than the longest period they try to extract — will perform very poorly at small frequencies. This implies that to select accurately periods equal to or smaller than 8 years from quarterly data, it is necessary to use *at least* a 32 points filter ($K = 16$). Thus, the low limit value of $K \geq 12$ suggested in [1] in the case of a passband of $[1.5, 8]$ yr is definitely too low.

Furthermore, Baxter and King argue that a good filter must not depend on the amount of data available, because this would imply a new computation of all the filter coefficients each time new data become available. Thus, time domain filtering should be preferred to frequency domain filtering. In our opinion it is not wise to neglect the new information that becomes available when N increases; as more information is added, it is crucial to take it into account to improve the quality of the filtered signal. This is particularly true in the case of short-length time series. It is exactly to stress the importance of this argument that we also plot the frequency responses of all the filter considered for 128 points (Figure 4, bottom panel): while the BK filter and the fixed-length symmetric CF filter (see below) remain unchanged, the resolution of the HP and of the HW filters is significantly improved and they become more accurate as new data are incorporated.

Finally, it is worth noticing that the same “right” behavior (Equation (8)) at the origin would have been ensured by the truncated filter (2) with no additional constraint but the requirement of harmonic cutoff frequencies, i.e. cutoff frequencies $\nu_{\{l,h\}}$ which are both chosen to be integer multiples of T^{-1} . Actually, it is easy to check by frequency response inspection that, apart from the constraint on the values of $\nu_{\{l,h\}}$, this much simpler filter performs better in the higher part of the spectrum than the Baxter-King filter.

3.3 The Christiano-Fitzgerald Filter

Christiano and Fitzgerald [3] build a filter using two new ingredients: (i) take into account the assumed spectral density of the original data; (ii) drop the conditions of stationarity and symmetry on the filter coefficients.

If the exact spectral density of the original data $U^{\text{exact}}(\nu)$ is known beforehand, the set of

coefficients $\{h_j\}$ is given by the solution of the optimization problem

$$\min_{\{h_j\}} \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} d\nu \left| \sum_{\{j\}} (h_j - h_j^{\text{ideal}}) e^{-i2\pi j\nu\Delta t} \right|^2 |U^{\text{exact}}(\nu)|^2, \quad (22)$$

which is equivalent to the minimization

$$|v_j - v_j^{\text{ideal}}|^2 = \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} d\nu |H(\nu) - H^{\text{ideal}}(\nu)|^2 |U^{\text{exact}}(\nu)|^2, \quad (23)$$

of the discrepancy between the ideally filtered data and the effectively filtered ones⁵.

According to different types of optimization problems, which give rise to different filters, the set of indexes $\{j\}$ could be constant symmetric $j = -K, \dots, K$, constant asymmetric $j = -K, \dots, K'$ or even a time-varying general one like $j = -(N - j), \dots, j - 1$. To obtain explicit solutions, Christiano and Fitzgerald assume different spectral density shapes. For instance, if $|U^{\text{exact}}(\nu)|^2$ is chosen as independent of frequency (white noise, referred to as *IID case* in [3]), the solution is simply given by truncating the ideal filter coefficients (1). If u_j has one unit root and $|U^{\text{exact}}(\nu)|^2$ goes like ν^{-2} for small frequencies but tends to a constant at large frequencies (the *near-IID case* in [3]), it is shown that the optimal coefficients are again obtained by truncating the ideal ones, but then subtracting from each coefficient the same constant to make their sum equal to zero and cancel the unit root. For a constant symmetric set of indexes, this gives the Baxter-King filter.

In the end, what is called the Christiano-Fitzgerald filter is obtained by taking the power spectral density $|U^{\text{exact}}(\nu)|^2 \propto \nu^{-2}$ for all frequencies, which is the case for a random walk process. The coefficients can be obtained explicitly and are given by truncating the ideal filter ones and then adjusting only h_{-K} and $h_{K'}$. In this way the sum of the left coefficients ($j = -K, \dots, 0$) and the sum of the right coefficients ($j = 0, \dots, K'$) are both zero and Equation (8) is satisfied. The filtering operation is the following:

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_{N-1} \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} \hat{h} & h_1 & h_2 & \dots & h_{N-3} & h_{N-2} & h_{N-1} \\ \hat{h} - h_0 & h_0 & h_1 & \dots & h_{N-4} & h_{N-3} & \hat{h} - h_{\{0, N-3\}} \\ \hat{h} - h_{\{0, 1\}} & h_1 & h_0 & \dots & h_{N-5} & h_{N-4} & \hat{h} - h_{\{0, N-4\}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hat{h} - h_{\{0, N-4\}} & h_{N-4} & h_{N-5} & \dots & h_0 & h_1 & \hat{h} - h_{\{0, 1\}} \\ \hat{h} - h_{\{0, N-3\}} & h_{N-3} & h_{N-4} & \dots & h_1 & h_0 & \hat{h} - h_0 \\ h_{N-1} & h_{N-2} & h_{N-3} & \dots & h_2 & h_1 & \hat{h} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \\ y_N \end{bmatrix}, \quad (24)$$

where the h_j 's are the ideal filter coefficients (2), $\hat{h} = h_0/2$, and we define $h_{\{0, j\}} = h_0 + h_1 + \dots + h_j$ to simplify the notation. It is evident — as the authors themselves stress⁶ — that this filter acts differently on each date, so that we actually have N different filters, represented by the columns of the matrix in (24). Another way to say it, is that the CF filter is time-varying, thus nonstationary. Moreover, at each time, the coefficients are asymmetric with respect to

⁵The problem (22) differs from the Baxter and King problem (20) by the presence of the weighting factor $|U^{\text{exact}}(\nu)|^2$.

⁶See p. 442 in [3].

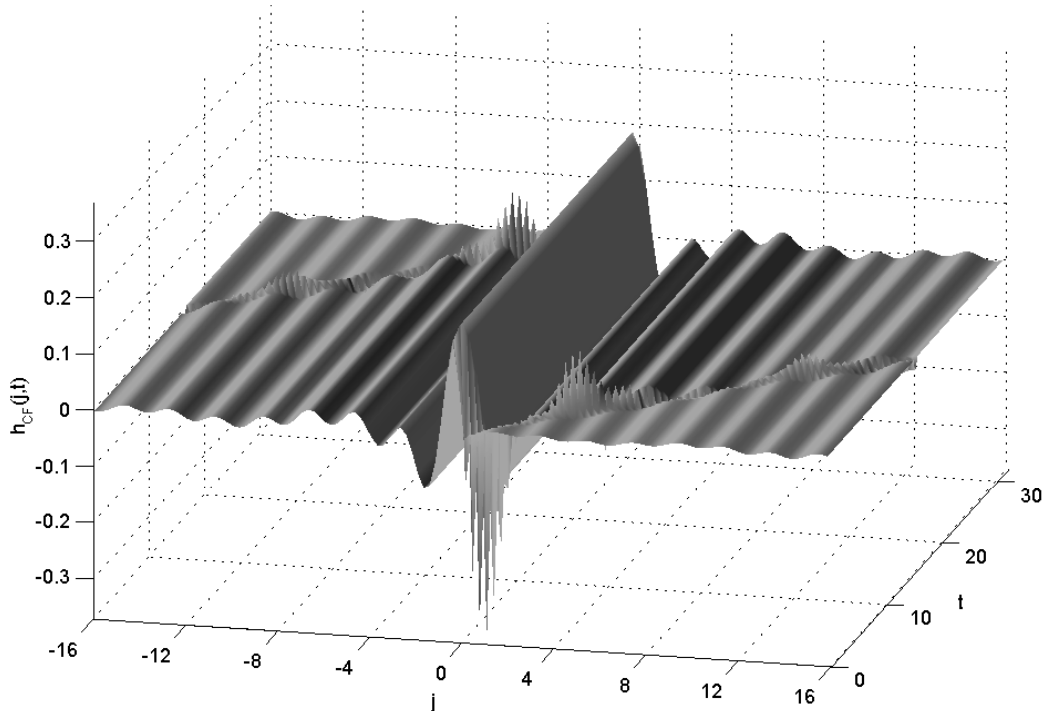


Figure 5: **CF Random Walk Filter (I): Coefficients.** This figure shows the time coefficients of a random walk CF filter selecting periodicities between 8 yr and 2 yr. It has been obtained with 128 points of quarterly data.

past and future data⁷, as shown in Figure 5. The effect of asymmetry is that the CF filter response is complex, thus it has a nonzero phase. The effect of nonstationarity is that the phase, as well as the real part of the frequency response, will not only depend on frequency but also on time.

The real and imaginary part of the CF filter frequency response and its standardized phase⁸

$$\tilde{\Phi}_{CF}(\nu, t) = \frac{\Phi_{CF}(\nu, t)}{2\pi\nu} = \frac{1}{2\pi\nu} \arctan\left(\frac{\Im H_{CF}(\nu, t)}{\Re H_{CF}(\nu, t)}\right) \quad (25)$$

are reported in Figures 6, 7 and 8, respectively. The figures have been obtained for 128 points of quarterly data and a passband of [2, 8] yr. By looking at the real part (Figure 6), we notice some leakage in the upper part of the spectrum, that is particularly pronounced for at least the first and last three years of data, which should thus be discarded. As for the phase, Figure 8 shows the spurious shifts induced in the signal by the CF filter. The absolute value of the phase reaches a maximum of approximately 1.6 quarters for certain components inside the passband. Some components can thus experience shifts up to ± 5 months, causing a maximum *relative* shift of almost one year. With the introduction of a spurious time- and frequency-dependent phase, timing and correlation properties among different frequency components within the series are therefore irreversibly modified and cannot be recovered. The same happens to correlation and causality relation among different series.

⁷Considering of course periodic boundary conditions, i.e. $y_{N+j} = y_{j \bmod N}$.

⁸The standardized phase measures the number of lags in units of Δt .

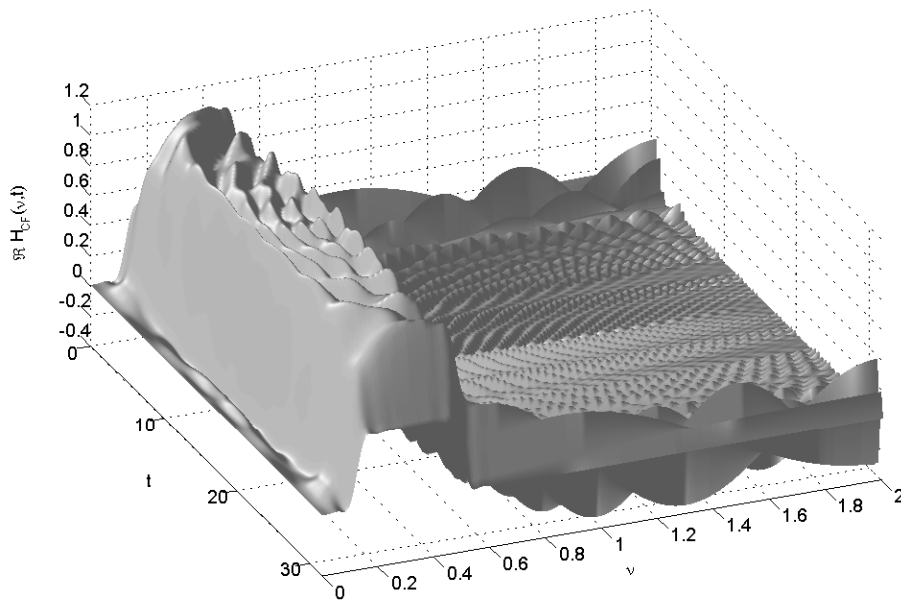


Figure 6: **CF Random Walk Filter (II): Real Part.** This figure shows the real part of the response function of the random walk CF filter. The parameters are the same as those of Figure 5.

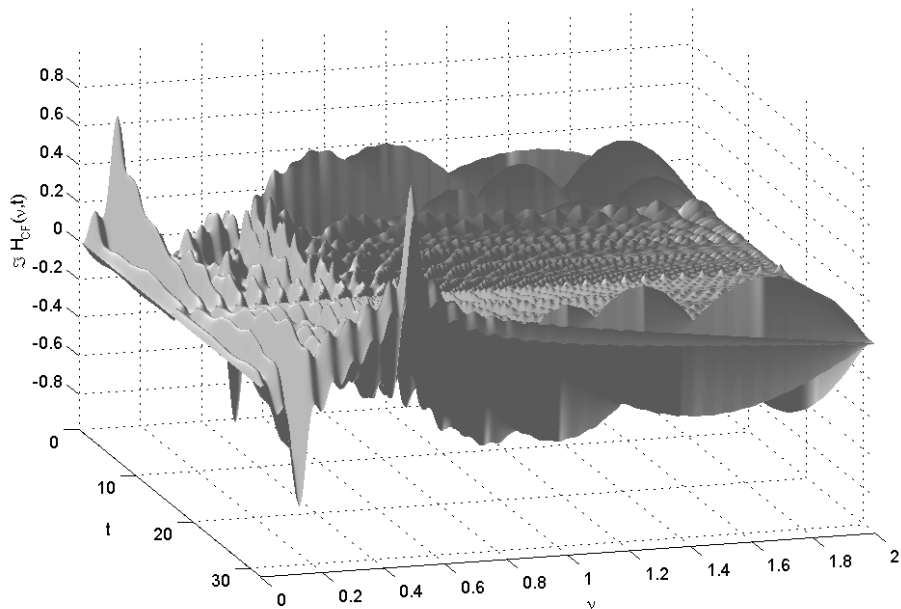


Figure 7: **CF Random Walk Filter (III): Imaginary Part.** This figure shows the imaginary part of the CF filter response function. Recall that, while the real part of the frequency response is an even function, this is an odd function with respect to $\nu = 0$. The parameters are the same as those of Figure 5.

In the aim of making a direct comparison with the “two-dimensional” filters, we also plotted the frequency response of the symmetric fixed-length version of the Christiano-Fitzgerald

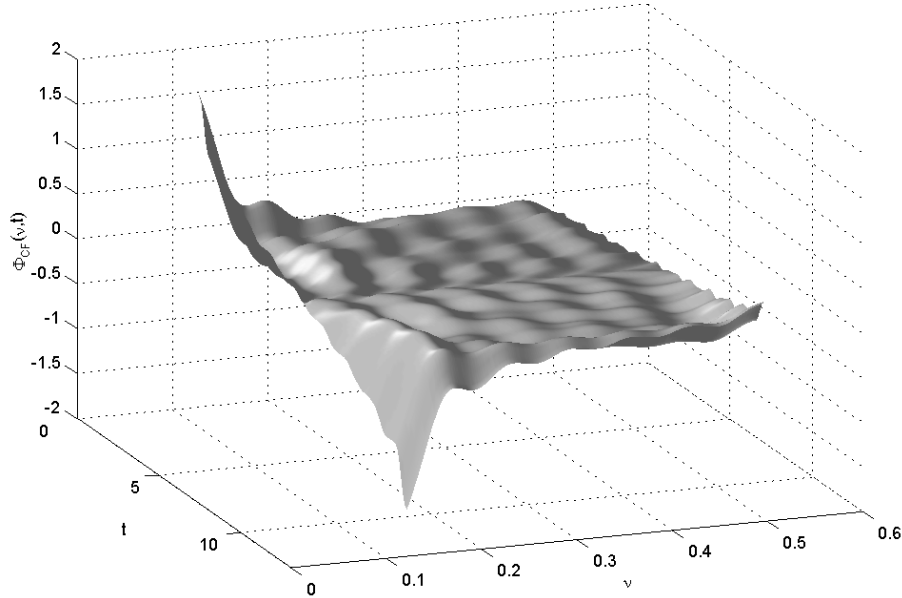


Figure 8: **CF Random Walk Filter (IV): Phase.** This figure shows the phase of the CF filter response function. We show only the portion lying in the passband $[2, 8]$ yr. The parameters are the same as those of Figure 5.

random walk filter⁹ in Figure 4. Because of the divergence of the random walk spectral density at small frequencies, the optimization puts a very strong emphasis on the behavior near the origin and attenuates very well these components (see Figures 4 and 6). This comes at the price of a bad removal of high frequency components. As in the case of BK filter, the poor performance at high frequencies stems from the increased discontinuity at the edges of the filter, once the spurious nonzero amplitude in the origin $H(0)$ has been transferred there. Of course, if the input signal has very small components at high frequencies, as is indeed the case of a real random walk, their leaking in the passband is irrelevant. On the other hand, if it is not sure whether the data can be really modeled as a random walk, as in the case of, e.g. a growth rate this filter is a rather poor approximation of the ideal one.

Going back to the CF random walk filter, in our opinion its advantages are minimal compared to the shortcomings. First of all, the introduction of the assumed spectral density in (22) to find the *optimal* filter adds a step in the computation of the filtered series (and not an easy one), i.e. the estimation of the data generating model. This is a radical change from the “philosophy” of the other three filters: from a purely descriptive tool to a model-based methodology. But this would be the least problem: there is a whole stream of research on model-based filtering (see e.g. [5] and [7]). In fact, the filter finally proposed by Christiano and Fitzgerald is not the optimal one, but the *nearly optimal*, i.e. the one they obtain by keeping the hypothesis of random walk for any time series. This procedure is dubious at least for two reasons: (i) the hypothesis plausibility, since, as they write themselves “This approach uses the approximation that is optimal under the (*in many cases false*) assumption that the data are generated by a pure random walk” (in [3], p.436, emphasis added); (ii) if all the beauty

⁹Notice that the frequency response plots in [3] are also obtained for the fixed-length symmetric random walk filter.

and improvement of the method was in the research of the optimal filter for a given series, why would one spoil it by always choosing the same spectral density, irrespective of the series? What is the point of using such a complex method to obtain a filter that might work ex post but for obscure (at least to us) reasons?

In a few words, besides the standard leakage and compression problems, this filter has two serious shortcomings: it is time-varying and asymmetric. This has at least two consequences. The first is that nothing can be said about the stationarity of the output signal, even if the input was itself stationary. The second is that the time and frequency dependent phase shift implies the loss of all timing relations between two series, a loss that can be crucial, as in the case of, e.g., the Phillips curve. A fair amount of work has been done on the issue of its “reappearance” in business cycle components (see, e.g. [6, 4] and [15, 3]), but one must be aware of the fact that once a phase shift is introduced like that of the CF filter, it changes the correlation function between inflation and unemployment, this may modify the form of the Phillips curve, making meaningless any further investigation.

Finally, we remark that nonstationary filters do not preserve pure harmonic signals, as we will show in Section 4.1.

4 Applications

4.1 Comparison on the Basis of an Artificial Series

To test the performance of all the considered filters on a known ground, we applied them to the artificial series given by a simple harmonic, thus stationary, signal

$$y_j = \sin\left(\frac{2\pi j}{24}\right) - 0.15 \sin\left(\frac{2\pi j}{6}\right), \quad (26)$$

where $j = 1, \dots, 120$ and the data are supposed quarterly, i.e. 30yr of data, with cutoff frequencies of 6 yr and 1.5 yr both submultiples of the signal duration.

The results are shown in Figure 9. The first thing we notice is that the HW filtered series suffer less from compression than the others, while the HP filtered series loses approximately the 50% of the original amplitude. Secondly, not considering for obvious reasons the BK filtered series, the edges of the HW series follow remarkably well those of the original series, as expected in the case of an harmonic series, while both the CF filtered and the HP filtered series depart from them. Thirdly, the CF filtered series has a progressive phase drift which affects differently each frequency component as the shape of the original signal is not preserved. Thus, not surprisingly, a stationary signal containing 1.5 yr and 6 yr components is turned, after application of the CF filter, into a nonstationary signal, the phase between the two components varying with time, while their *mean* frequencies has been increased to $(1.49 \text{ yr})^{-1}$ and $(5.7 \text{ yr})^{-1}$ respectively.

All these shortcomings become less important when the series length is

increased, but operationally the CF filter rather than a filter in the acknowledged sense, is best described as a smoothing procedure whose effect on frequency components can hardly be established.

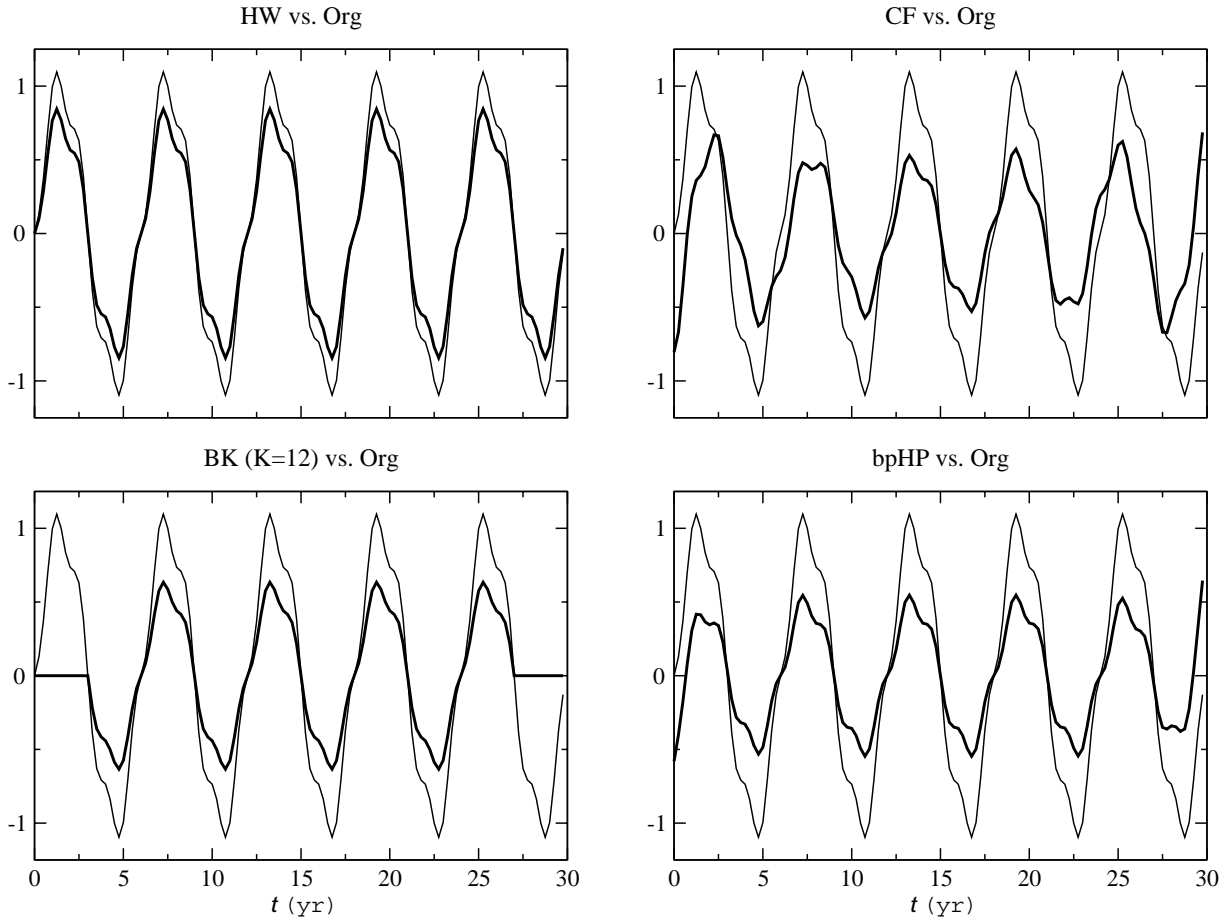


Figure 9: **Application to an Artificial Series.** A 120-points quarterly data signal containing 1.5 yr and 6 yr components (thin line, all panels) is filtered in the band [1.5 yr, 6 yr] using the HW filter (thick line, top-left panel), the CF random walk filter (thick line, top-right panel), the BK filter with $K = 16$ (thick line, bottom-left panel) and the bandpass HP filter ($\lambda_{low} = 215.3225$ and $\lambda_{low} = 1$, thick line, bottom-right panel).

4.2 Application to the Euro Zone Gross Domestic Product

In this section, we show the effect of different filters on real data. We chose as an example the quarterly series of Euro zone gross domestic product (Figure 10, left panel), from 1970:I to 2001:IV, thus 32 years (128 data points). The chosen band was [2, 8] yr, thus the frequency responses of the applied filters are exactly those plotted in Figure 4 and Figures 6 to 8. We emphasize that we chose the above-mentioned band since it is virtually equivalent to the “definition” of the business cycles ([1.5, 8] yr, see, e.g., [1]) and the duration of the series is multiple of both cutoff frequencies, which are thus harmonic with respect to the signal. The length of the series also was modified at this purpose. The condition of harmonicity of the cutoff frequencies with respect to the signal duration ensures the best performance of frequency filters.

We chose the CF filter random walk filter, in its recommended non-symmetric non-sta-

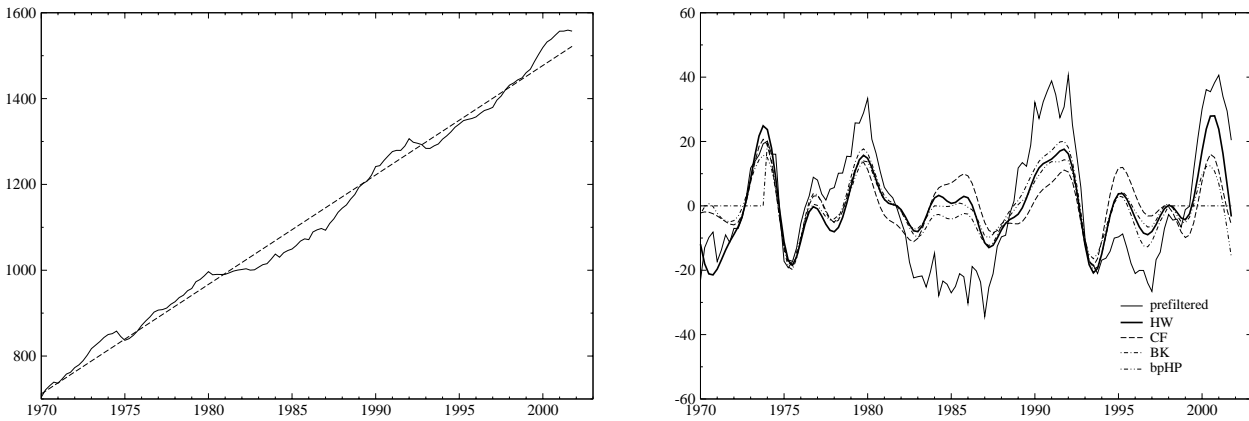


Figure 10: **Raw and Filtered Euro Zone GDP series.** The original series is shown (left panel) together with its least square fit. In the right panel we show the Hamming-windowed (full line), the Christiano-Fitzgerald (dashed line), the Baxter-King (dot-dashed line) and the Hodrick-Prescott (dot-dot-dashed) bandpass filtered series plotted against the detrended prefiltered Euro Zone GDP series.

tionary version. For the BK filter we choose $K = 16$ instead of $K = 12$ as suggested in [1] for the reasons we stated in Section 3.2. As for the HP filter, we applied it twice, once with $\lambda \approx 677$ (677.1298) to obtain a 8 yr-cutoff highpass filter and the other with $\lambda \approx 3$ (2.9142) for the 2 yr-cutoff highpass filter. We subtract the second series from the first, so that we finally select the desired frequency band from the original series. We point out that for both the HW and the CF filters, there is an operation of linear detrending of the data prior to that of filtering. The least squares line is shown in Figure 10 (left panel). The R^2 coefficient for the regression is 0.99 with $t_{Student} \approx 104$, thus highly significant. The detrended series was then *prefiltered* by means of a lowpass HW filter with $\nu_L = (16 \text{ yr})^{-1}$, to ensure covariance stationarity and consequently making possible the definition of the spectrum (dotted line in Figure 11). To the resulting series we finally applied the four bandpass filters.

The filtered series are shown in Figure 10 (right panel). Despite the fact that they all have similar shapes, the amplitude of their fluctuations is different, especially near the edges. In particular, in the cases of the HP, the BK and the CF filters, it is always smaller than for the HW filter. In the HP case, this is due to the strong filter compression, whereas in the BK case, it is an evidence of both the compression and the truncation of the coefficients. Instead, it reflects the non-stationarity of the CF filter, whose response amplitude decreases close to the ends of the data set (see Figure 6).

To assess the quality of the filtering procedures, we decided to plot, as an estimate of the true spectra, their periodograms

$$P_u(k) = \Delta t \sum_{J=-(N-1)}^{N-1} \gamma_{uu}(J) e^{-i2\pi Jk/N}$$

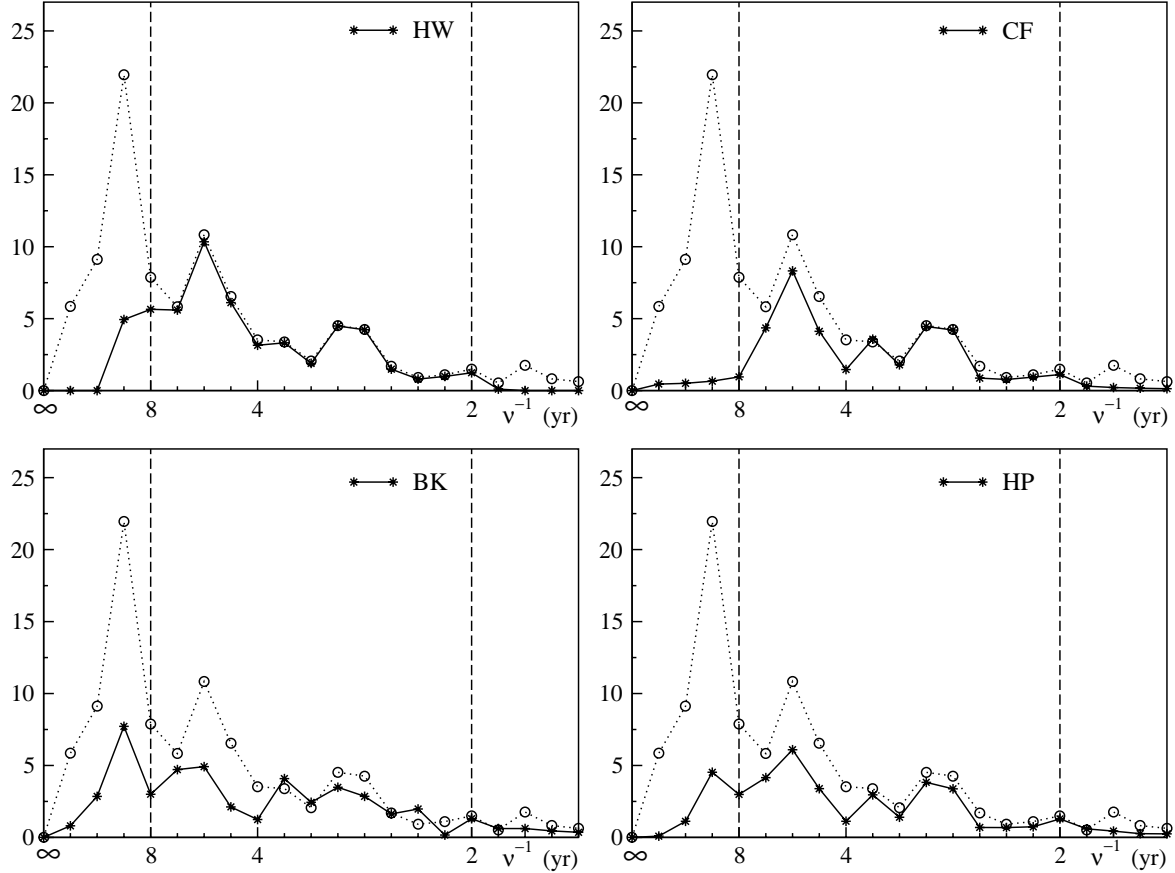


Figure 11: **Spectra Obtained by the Different Filtering Operations.** The spectra obtained by means of the four different filters (full line, all panels) are plotted against the spectrum of the detrended prefiltered Euro zone GDP. The corresponding series are plotted in Figure 10.

$$= \Delta t \sum_{J=-(N-1)}^{N-1} \gamma_{uu}(J) \cos \frac{2\pi Jk}{N}, \quad (27)$$

where $\gamma_{uu}(J) = \gamma_{uu}(-J) = N^{-1} \sum_{j=-(N-J)}^{N-J} (u(j) - \bar{u})(u(j+J) - \bar{u})$ is the estimator of the series autocovariance function at lag J . Of course, the objection of non consistency of the chosen spectral estimator may be raised. Nevertheless, each filter actually acts on the periodogram; besides the comparison is made at constant N . For these reasons we believe that the periodogram can be used to check which, among the four filters, is the best approximation to the ideal filter.

The periodograms of the four filtered series are plotted against that of the prefiltered in Figure 11¹⁰. The HW filtered series spectrum (top-left panel) almost exactly reproduces the original frequency components inside the selected band and vanishes outside, except for the frequency components near the band edges, which are compressed by the (necessary) smoothing

¹⁰Following the discussion in Section 3.3, we plotted the CF filtered series spectrum assuming its stationarity, since the spectrum (periodogram) is only definable for stationary series.

of the rectangular window. The CF filter causes significant compression (top right panel) in the left part of the band. Moreover, though it performs a better attenuation than the HW filter near the lower bound of the frequency band, there is a bothering leakage left at low frequencies (high periods), between $\nu = \nu_l$ and $\nu = 0$. Apart from compression, which is however more pronounced than in the case of the CF filter, the BK filter (bottom left panel) does not preserve the relative amplitudes of the components. This is due to the fluctuation of its frequency response in the passband. Furthermore, the BK filtered series spectrum shows high leakage in the low frequency stopband. Contrary to the previous one, the HP filter, instead, leaves the proportion among components unaffected inside the passband. Furthermore, it has a much lower leakage at low frequencies compared to the BK filter. This was visible also in Figure 4 (bottom panel).

5 Conclusions

To obtain an “ideal” filter, i.e. a filter selecting a finite range of frequencies with infinite resolution, we should have an infinite number of data. If the number of data is finite — which is like having a lag window on an infinite-length signal — an ideal filter cannot be realized and compromise is necessary. The simple truncation of the filter coefficients to match the signal length produces a filter which is optimal in the least-squares sense, but displays strong leakage and large ripples in its frequency response. Most often, one is prepared to give up a faster transition between passband and stopband to obtain a more reduced leakage. This is the aim of the windowed filters we are proposing here. In particular, the HW filter attenuates undesired components by a factor of more than 100, even for short-length time series.

The windowed filters can be designed for a given signal length and used to filter either in time or in frequency domain. We preferred the latter, since using the whole signal length to compute filtered values, improves the frequency resolution by exploiting all available information. The resulting filters are both stationary and symmetric, which are fundamental properties to preserve *all* timing relations between frequency components within the same series or across different series. Moreover, bandpass and highpass windowed filters are able to stationarize at least $I(2)$ processes.

The good performance of the HW filter in rejecting the off-band frequency components has been checked by means of a comparison with the BK, the CF and the HP filters. We presented a critical and in-depth review of these last three filters and confronted them to our HW filter. This was done on the basis of their frequency response and their action on both artificial and macroeconomic time series. The HW filter proves to be a much better performing tool for the empirical study of business cycle and for establishing the correlation properties of the variables of interest.

References

- [1] Baxter, M. and R. G. King, *Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series*, The Review of Economics and Statistics, November 1999, 81(4): 575-593, and NBER Working Paper 5022 (February 1995).

- [2] Benati, L., *Band-Pass Filtering, Cointegration and Business Cycle Analysis*, Bank of England, working paper (2001).
- [3] Christiano, L. and T. J. Fitzgerald, *The Band-Pass Filter*, International Economic Review, vol. 44, no. 2, May 2003, and NBER Working Paper 7257 (July 1999).
- [4] Gaffard, J. L. and A. Iacobucci, *The Phillips Curve: Old Theories and New Statistics*, in “Industrial Dynamics and Employment in Europe”, European Commission Scientific Report, SOE1-CT97-1055 (2001).
- [5] Gomez, V., *The Use of Butterworth Filters for Trend and Cycle Estimation in Economic Time Series*, Journal of Business & Economic Statistics, July 2001, Vol. 19, No. 3.
- [6] Haldane, A. and D. Quah, *UK Phillips Curves and Monetary Policy*, Journal of Monetary Economics, Special Issue: “The Return of the Phillips Curve”, vol. 44, no. 2, (1999), pp. 259-278.
- [7] Harvey, A. C. and T. Trimbur, *General Model-Based Filters for Extracting Cycles and Trends in Economic Time Series*, The Review of Economics and Statistics, May 2003, 85(2): 244-255.
- [8] Harvey, A. C. and T. Trimbur, *Trend Estimation, Signal-Noise Ratios and the Frequency of Observations*, 4th EUROSTAT Colloquium on Modern Tools for Business Cycle Analysis, Luxembourg, October 2003.
- [9] Hodrick, R. J. and E. C. Prescott, *Postwar US Business Cycles: An Empirical Investigation*, Journal of Money, Credit, and Banking, vol. 29-1 (1997), pp. 1–16.
- [10] Jenkins, G. M. and D. G. Watts, *Spectral Analysis and Its Applications*, Holden-Day, San Francisco, (1969).
- [11] King, R. G. and S. T. Rebelo, *Low frequency filtering and real business cycles*, Journal of Economic Dynamics and Control, 17 (1993), pp. 207–231.
- [12] Murray, C. J., *Cyclical Properties of Baxter-King Filtered Time Series*, The Review of Economics and Statistics, May 2003, 85(2): 472-476.
- [13] Oppenheim, A. V. and R. W. Schaffer, *Discrete-Time Signal Processing*, second edition, Prentice-Hall, New Jersey, 1999.
- [14] Ravn, M. O. and H. Uhlig, *On Adjusting the HP-Filter for the Frequency of Observations*, The Review of Economic and Statistics, May 2002, 84(2): 371–380.
- [15] Stock, J. H. and M. W. Watson, *Business Cycle Fluctuations in US. Macroeconomic Time Series*, in “Handbook of Macroeconomics”, vol. 1a, J. Taylor and M. Woodford eds., North-Holland, Amsterdam (1999).