Computational Economics (2005) 25: 75–102
DOI: 10.1007/s10614-005-6276-7 

# A Frequency Selective Filter for Short-Length Time Series

ALESSANDRA IACOBUCCI[1,2,*] and ALAIN NOULLEZ[3]
[1] *OFCE, 69, quai d'Orsay, F-75340 Paris, Cedex 07, France;*
*E-mail: alessandra.iacobucci@sciences-po.fr*
[2] *GREDEG, 250 rue Albert Einstein, F-06560 Valbonne, France;*
[3] *Observatoire de Nice, B.P. 4229, F-06304 Nice Cedex 4, France*

Accepted 25 April 2005

**Abstract.** An effective and easy-to-implement frequency filter is proposed, obtained by convolving a raised-cosine window with the ideal rectangular filter response function. Three other filters, Hodrick–Prescott, Baxter–King, and Christiano–Fitzgerald, are thoroughly reviewed. A bandpass version of the Hodrick–Prescott filter is also introduced and used. The behavior of the windowed filter is compared to the others through their frequency responses and by applying them to both quarterly and monthly artificial, known-structure series and real macroeconomic data. The windowed filter has almost no leakage and is better than the others at eliminating high-frequency components. Its response in the passband is significantly flatter, and its behavior at low frequencies ensures a better removal of undesired long-term components. These improvements are particularly evident when working with short-length time series, which are common in macroeconomics. The proposed filter is stationary and symmetric, therefore, it induces no phase-shift. It uses all the information contained in the input data and stationarizes series integrated up to order two. It thus proves to be a good candidate for extracting frequency-defined series components.

**Key words:** frequency domain filtering, spectral methods, HP filter, Baxter–King and Christiano–Fitzgerald bandpass filters, business cycles

**JEL codes:** C10; C14; E32

## 1. Introduction

There are several ways for 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 harmonic components, each having a time-invariant amplitude 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 and scale description parameter respectively – are also an alternative, especially in the case of nonstationary or intermittent signals. These techniques allow a detailed insight into the data structure. However, they are not easily implemented with short-length series and are not broached in this paper.

A 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 a filter that preserves all frequencies in a given range and completely removes those outside it (the so called *ideal filter*). Indeed, especially in the case of short-length series, abrupt variations in the frequency response give rise to the *Gibbs phenomenon*, the appearance of spurious artificial fluctuations in the filtered signal (see, among others Hamming, 1998; Oppenheim and Schafer, 1999; Priestley, 1981). Therefore, it is important to design good approximations to the ideal filter – "good" relative to some optimization criteria, like the (weighted) difference between the desired and the effective response.

We propose here the Hamming/Hanning-windowed (HW) filter, which provides a simple and efficient solution to an ideal filter approximation. It is obtained by *windowing*, which is a well-known technique in engineering (Oppenheim and Schafer, 1999). It consists in smearing the ideal filter response with a lag window and it leads to good attenuation of the spectral power outside the passband, allowing almost complete removal of undesired frequency components. Its only drawback is a negligible widening of the transition between the selected band (*passband*) and the remaining spectral components (*stopband*), which implies that a negligible part of the frequency components lying near the edge of the chosen band may be present in the filtered series.

To show the qualities and improvements of the HW filter, we compare it to those most widely used in macroeconomics for trend and cycle extraction, namely the filters in Hodrick and Prescott (1997), Baxter and King (1999) and Christiano and Fitzgerald (2003). We thoroughly review these filters and, in particular, we visualize the time coefficients, gain, and phase of the Christiano–Fitzgerald filter, plotting them as surfaces in three-dimensional space.

In the next section, we briefly recall the issues of optimal filter design. In the third section, the HW filter is introduced, together with a sketch of its computing algorithm. The fourth section contains an extensive critical overview of the the Hodrick–Prescott (HP), the Baxter–King (BK), and the Christiano–Fitzgerald (CF) filters, and a detailed comparison of their frequency responses to that of the HW filter. The fifth section is devoted to applications, that allow a thorough comparison of the four considered filters on practical grounds. We first analyze the effect of each filter on the spectral density of a real series, the quarterly EU zone GDP, to visualize leakage and compression effects. We then apply the four filters to a simple artificial series, to show more clearly their effects on individal frequencies and to highlight the phase-shift induced by the HP and CF filters. Finally, we show how some filters might affect timing and causality relations among series by studying crosscorrelation properties, first of two simple artificial series, and then between US unemployment rate and annualized inflation monthly data. The sixth section concludes.

## 2. The Optimal Filtering Problem

It is known that an infinite series would be required to obtain the ideal (rectangular) frequency selective filter $H^{\text{ideal}}(\nu) = \Theta(\nu - \nu_l)[1 - \Theta(\nu - \nu_h)]$, where $\Theta(x)$ is the Heaviside step function and $\nu_h$ and $\nu_l$ represent the high and low cutoff frequency, respectively. The aim of optimal filtering theory consists in finding the best approximation to $H^{\text{ideal}}(\nu)$ in the case of finite series.

We deem the essential requirements to be met in solving the optimal filtering problem to be: (i) the approximated filter should leave as much information unaffected as possible over the specified range of periodicities it is supposed to extract; (ii) it should not introduce spurious phase shifts, which would imply a modification of timing relations among different series or among different frequency components within the same series; (iii) its output must be stationary. (Hereafter we use lower case variables to denote time functions, and upper case variables for their corresponding discrete Fourier transform.)

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 $\Delta t$ the sampling periodicity:

$$
\begin{aligned}
v_j = h_j \circledast u_j &\equiv \sum_{n=0}^{N-1} h_n\, u_{j-n \bmod N} \\
&= \frac{1}{N} \sum_{k=-N/2}^{(N-1)/2} H_k U_k e^{i2\pi jk/N} .
\end{aligned}
\tag{1}
$$

Since we are dealing with a finite discrete series, the frequency $\nu_k = k/(N\Delta t)$ and the time $t_j = j\Delta t$ variables are also discrete and indexed by $k$ and $j$, respectively. Moreover, because of discretization, the absolute value of $\nu_k$ is bounded by $\nu_{\text{Nyq}} = (2\Delta t)^{-1}$, the so-called *Nyquist frequency*. It is easy to see from (1) that the filter $H_k$ is not causal, since each point of the output series is computed by means of past as well as future values of the input series.

From the expression of the filter frequency response

$$
H_k = \Delta t \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} h_n e^{-i2\pi nk/N} , \tag{2}
$$

we deduce that a filter that is real in the time domain ($h_j = h_j^*$) is symmetric in the frequency domain ($H_k = H_{-k}$), and *vice versa*. Therefore, if we want real signals to remain real after filtering, both time *and* frequency response functions have to be real and symmetric, to avoid time and/or frequency shifts. Indeed, it is easy to see that, if the filter $H_k$ is a complex function, different frequencies undergo different phase shifts and the timing relations among components are destroyed.

Equation (2) implies that the sum of the windowed filter coefficients $h_j$ is set by the frequency response at the origin

$$H_0 = \Delta t \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} h_n. \tag{3}$$

This allows the removal of the signal mean in the case of lowpass and bandpass filters,

$$H_k|_{k=0} = 0. \tag{4}$$

If we add the following condition on the filter first time derivative:

$$\Delta H_k|_{k=0} = 0, \tag{5}$$

or its time domain equivalent

$$\sum_{n=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} n h_n = 0, \tag{6}$$

then the elimination of two unit roots is ensured.

The circular convolution (⊛) in (1) implies that the finite signal is replaced by its periodic version $u_{N+j} = u_{j \bmod N}$, the maximum period length being implicitly assumed by the Fourier transform to be $T = N\Delta t$. This imposed $T$-periodicity is in general false and affects the analysis. Indeed, it introduces an artificial discontinuity at the edges of the time series and causes the Gibbs effect. As is shown in the next section, these oscillations are due to the form of the discrete Fourier transform of a rectangular window of width $T = N\Delta t$ (Figure 1, dotted line).

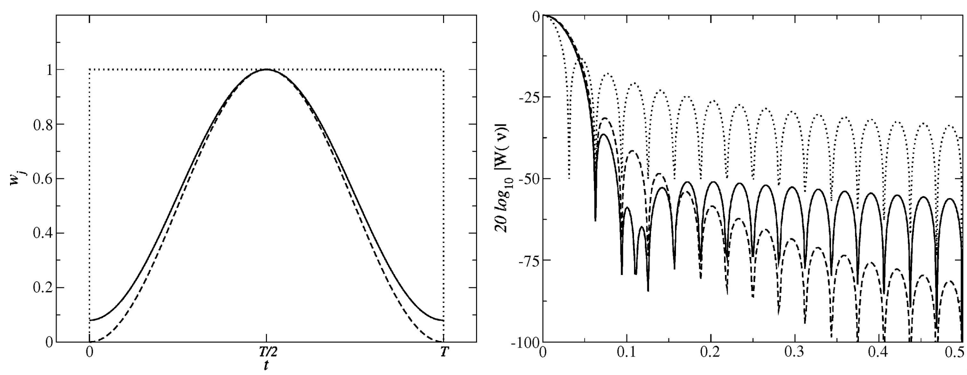


*Figure 1.* Window functions and their frequency responses: rectangular (dotted line), Hanning (dashed line), Hamming (full line).

There are three ways to minimize this effect. The first is to choose the length of the series $T$ as a multiple of the highest period that is likely to occur. This is a very difficult task, since it requires knowledge of the data-generating process and would, in any case, imply some data loss. The second is to choose both cutoff frequencies as integer multiples of $T^{-1}$, which is easier to do, but might also imply data loss. The third, if the series has a deterministic trend, for instance a polynomial function of time, is to subtract the least-squares fit *before* filtering. In fact, the operation of detrending and that of filtering do not commute, since a polynomial and a periodic term are not orthogonal. Subtracting the OLS regression line can always be done to reduce the edge discontinuities that appear if there are nonharmonic frequencies in the signal. The effect of this discontinuity is, however, reduced to negligible amplitudes by the use of windowing.

## 3. The Windowed Filter

The truncation to $N$ points of the ideal filter time coefficients

$$h_j^{\text{ideal}} = \frac{\sin(2\pi \nu_h j \Delta t) - \sin(2\pi \nu_l j \Delta t)}{\pi j}, \quad j = -\infty, \ldots, \infty,$$
$$h_0^{\text{ideal}} = \lim_{j \to 0} h_j^{\text{ideal}} = 2\Delta t(\nu_h - \nu_l), \tag{7}$$

where $j = -\infty, \ldots, \infty$, amounts to multiplying (7) by an $N$-wide rectangular lag window, yielding

$$h_j = \frac{\sin\left[\frac{2\pi j}{N}\left(k_h + \frac{1}{2}\right)\right] - \sin\left[\frac{2\pi j}{N}\left(k_l - \frac{1}{2}\right)\right]}{N \sin\left(\frac{\pi j}{N}\right)} \quad j = 1, \ldots, N-1,$$
$$h_0 = \lim_{j \to 0} h_j = \frac{2}{N}(k_h - k_l + 1). \tag{8}$$

The truncation induces fluctuations of large amplitude and slow decay in the response function – the Gibbs effect again – 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 to the rectangular lag 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 the signal processing literature (Jenkins and Watts, 2000; Priestley, 1981; Oppenheim and Schafer, 1999) and is not a univocally defined problem because different optimization criteria could be used, among which: (i) the minimization of the *leakage* *factor*, that is, the ratio of the power in the sidelobes to the total window power (see Figure 1); (ii) the minimization of the mainlobe width (the highest peak in Figure 1, right panel), which is proportional to the transition band width – the width of the band that arises between the stopband and the passband, as a consequence of the *transition* finiteness of real series; (iii) the minimization of the highest or the first sidelobe peak height. Notice that there is a trade-off between (i) and (ii): the lower the leakage factor, the wider the mainlobe width, and *vice versa*. This is shown in Figure 1: the rectangular window indeed has the narrower mainlobe width, which implies a steeper transition between the passband and the stopbands, but at the price of a high leakage factor.

From these and other possible criteria, a broad range of windows arise (triangular, Bartlett, Parzen, Blackman, Tukey–Hanning, Hamming, . . . ; see Priestley, 1981). We single out the *raised-cosine* (or *general Tukey*) windows

$$w_j^{\mathrm{Han}} = a - (1-a)\cos\left(\frac{2\pi j}{N}\right), \quad \frac{1}{2} \le a < 1, \tag{9}$$

which have spectral windows of the form

$$W(\nu) = \left[(2a-1) + \frac{2(1-a)}{1-(\nu T)^2}\right]\frac{\sin(\pi\nu T)}{\pi\nu T}. \tag{10}$$

More precisely, we consider the so-called Hanning ($a = 0.5$) and Hamming ($a = 0.54$) windows (the H-windows in the following), which represent the best compromise for our purposes in building a bandpass filter.

The Hanning window is continuous and vanishes at both zero and $N$ along with its first derivative (see Figure 1, left panel), while the Hamming window is not continuous at the edges of the interval and is obtained by a judicious combination of the Hanning and the rectangular windows. As for their frequency responses, the H-windows have a mainlobe width that is almost twice as wide as that of the rectangular window. This implies that the filters obtained by H-windowing have a transition band approximately twice as wide as that of the ideal filter – the price paid for smoothing. The Hanning spectral window performs better than the Hamming one in the upper part of the spectrum (high $\nu$), and this is useful with long time series with high frequency resolution, like those typical in finance. However, the Hamming spectral window attenuates the first sidelobes amplitude better, and hence, it is the appropriate window to use when frequencies close to the edges of the passband have to be eliminated – especially in the case of short-length series more common in macroeconomics.

Both the Hanning and Hamming spectral windows 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.

Obviously, different windows may perform better as regards one specific criterion. For instance, the Blackman and Parzen windows have a lower leakage factor and a better sidelobe attenuation than the H-windows, but a wider transition band, that affects the performance of the bandpass filter, especially in the case of short-length series. Not restraining ourselves to “static” windows, we could also consider other windows, like the Gaussian or the Kaiser windows, which have a parameter that allow to establish quantitatively and to modify the ratio between leakage factor and transition (mainlobe) width. Nevertheless, the use of such windows leads to much more complicated computations and a marginal gain with respect to H-windows.

### 3.1. THE HW FILTER ALGORITHM

We now introduce the proposed filtering procedure, described by the following algorithm:

(i) subtract, if needed, the least-square regression line to remove the artificial discontinuities introduced at the edges of the series by the Fourier transform;

(ii) compute the discrete Fourier transform of the signal $u_j$

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

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

$$V_k = (W_k * H_k)\, U_k = \sum_{k'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} W_{k'} H_{k-k'} U_k$$

$$= \left[ \left( \frac{(1-a)}{2} H_{k-1} + a\, H_k + \frac{(1-a)}{2} H_{k+1} \right] U_k, \quad k = 0, \ldots, \lfloor N/2 \rfloor,$$

where $H_k$ is defined by the frequency range as

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

and $W_k$ is the discrete window Fourier transform, which has only three nonzero components at frequencies 0 and $\pm(N\Delta t)^{-1}$;

(iv) compute the inverse transform

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

Windowing the filter response in the frequency domain by convolution of the ideal response with the spectral window, as in (iii), is computationally less expensive than time domain multiplication. In fact, as stressed above, both the Hamming and Hanning discrete spectral windows have only three nonzero components. Also

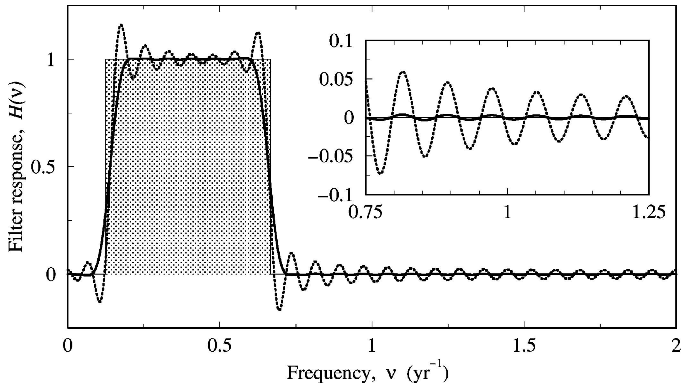


*Figure 2.* Windowed (full line) and truncated (dotted line) filters frequency response.

note that the HW filters satisfy conditions (4) and (5), and thus stationarize an $I(2)$ series.

From now on we deal only with the Hamming-windowed filter, since it is more suitable for short macroeconomic series. But the differences between the two filters in comparison with the others are almost imperceptible. Figure 2 shows the frequency 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 most of the undesired frequencies than it is to have a sharp discrimination between frequencies at the edges of the passband. Summing up, this filtering 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.

## 4. Overview of the Most Popular Bandpass Filters

In this section, we provide a critical review of some of the frequency-selective filters that are widely used in the literature for trend and cycle extraction, namely the HP filter (Hodrick and Prescott, 1997), the BK filter (Baxter and King, 1999) and the CF filter (Christiano and Fitzgerald, 2003). Some interesting by-products of this review are: the finding of an explicit formula for the computation of the HP filter $\lambda$-parameter as a function of the sampling periodicity and of the cutoff frequency; the consequent building of a bandpass version of the HP filter, in order to make a fairer comparison among all the filters (see Section 5); the formal expression of the minimum moving-average order of the BK filter (the parameter $K$), to obtain the appropriate resolution according to the cutoff frequencies values; and the three-dimensional plots of the CF filter time coefficients, gain and phase.

### 4.1. THE HODRICK–PRESCOTT FILTER

The Hodrick–Prescott cyclical component $v_j^{\mathrm{HP}}$ 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

$$\min_{\{g_j\}} \sum_{j=0}^{N-1} [(u_j - g_j)^2 + \lambda\,(g_{j+1} - 2\,g_j + g_{j-1})^2], \tag{11}$$

which minimizes the sum of the norm of the cyclical component $\|v_j^{\mathrm{HP}}\| = \|u_j - g_j\|$ 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 $\lambda$ 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 $\lambda$, the growth component $g_j$ tends to the OLS line calculated from the data.

The solution of (11) for $N \to \infty$ can be found explicitly in the frequency domain (King and Rebelo, 1993) and leads to the following expression for the frequency response function

$$\begin{aligned} H^{\mathrm{HP}}(\nu) &= \frac{4\,\lambda(1-\cos(2\pi\,\nu\Delta t))^2}{1+4\,\lambda(1-\cos(2\pi\,\nu\Delta t))^2} \\ &= \frac{16\,\lambda\ \sin^4(\pi\,\nu\Delta t)}{1+16\,\lambda\ \sin^4(\pi\,\nu\Delta t)}, \end{aligned} \tag{12}$$

where $G(\nu)$ is the Fourier transform of the series growth component $g_j$. From the previous 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

$$\nu_c = (\pi\,\Delta t)^{-1} \ \arcsin\left(\frac{\lambda^{-1/4}}{2}\right); \tag{13}$$

that is, $\nu_c = 0.0252\,(\Delta t)^{-1}$ when $\lambda = 1600$. Hence the HP filter, in the configuration suggested by the authors for quarterly data, selects periodicities shorter than approximately 10 years, but has the disadvantage of a wide transition band (see Figure 3). The frequency response goes as $\lambda(2\pi\,\nu\Delta t)^4$ at low frequencies; hence it behaves as a fourth-difference filter and can stationarize an $I(4)$ process.

A recurring issue when using the HP filter is the value of the parameter $\lambda$ to use when dealing with annual or monthly data. This has been studied in, among others, Ravn and Uhlig (2002), which finds

$$\lambda_s = s^4\,\lambda_q, \tag{14}$$

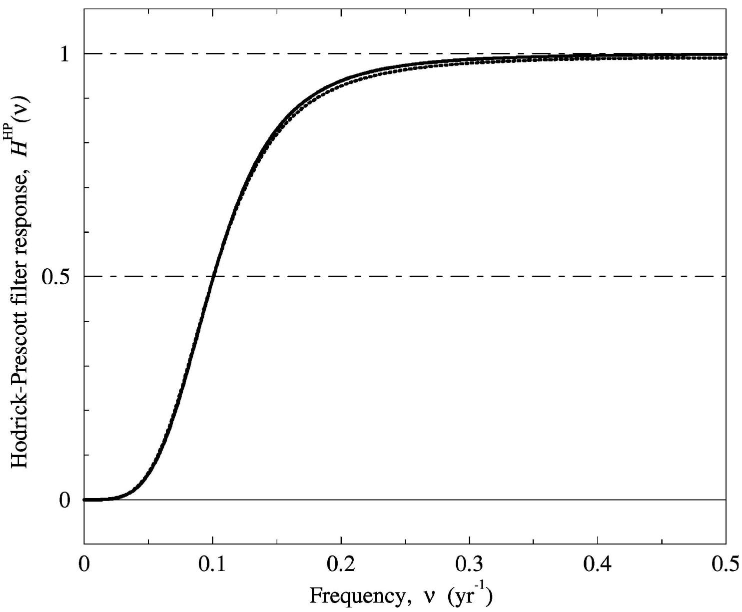


*Figure 3.* Hodrick–Prescott filter frequency response.

where $\lambda_q$ is the value of the parameter for quarterly data (the usual 1600), $s$ is the relative alternative sampling frequency (equal to $1/4$ for annual data and 3 for monthly data), and $\lambda_s$ the corresponding parameter value. As for the dependence of $\lambda$ on the cutoff frequency, by noticing that the HP filter belongs to the wider class of Butterworth filters, Gomez (2001) indicates the expression

$$\lambda = [2\sin(x_c/2)]^{-4} , \qquad (15)$$

where $x_c$ stands for the reduced angular frequency. A more general and comprehensive work is Harvey and Trimbur (2003b). This analyzes the dependence of $\lambda$ on the cutoff frequency and the sampling frequency in a model-based approach to filters. In particular, it examines how the variation of $\lambda$ can change the structure of an unobserved component model by modifying, for example, the correlation between components.

To find the explicit dependence of $\lambda$ on both the sampling frequency and the cutoff frequency $\nu_c$, we can directly solve the frequency response (12) for $\lambda$. Through (13), we obtain

$$\lambda = [2\ \sin(\pi \nu_c \Delta t)]^{-4}. \qquad (16)$$

Consequently, for a value of $\nu_c = (9.92\,\text{year})^{-1}$ such that $\lambda = 1600$ for $\Delta t = 4^{-1}$ years, we find that $\lambda = 6.68$ when $\Delta t = 1$ year (see Figure 3) and $\lambda = 129660$

when $\Delta t = 12^{-1}$ years. Equations (15) and (16) 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 for the reduced frequency $\nu_c \Delta t$ (large values of $\lambda$), the formula (14) is equivalent to (16) and of course to (15), and both are equivalent to

$$\lambda \approx (2\pi \nu_c \Delta t + O((\nu_c \Delta t)^3))^{-4}. \tag{17}$$

Therefore $\lambda$ varies as the inverse fourth power of $\nu_c$ or $\Delta t$, as can be guessed by inspection of (11), where $\lambda$ multiplies the square of the second difference of $g_j$.

By means of the formula (16) or (17), 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 $\lambda$ values, as shown in Gomez (2001) and Harvey and Trimbur (2003a) for Butterworth filters. Figure 4 shows the [2, 8] years bandpass HP filter obtained through (12) for a sample of $N = 128$ quarterly data points. Since the length of this filter, that is, the number of points involved in the calculation of one point of the filtered series, 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 quite a poor approximation to the ideal bandpass filter. Furthermore, since it is a recursive filter – the initial condition $g_{-1}$ is indeed required for the solution of the optimization problem (11) – its finite version is nonsymmetric and introduces a phase shift near the series edges.

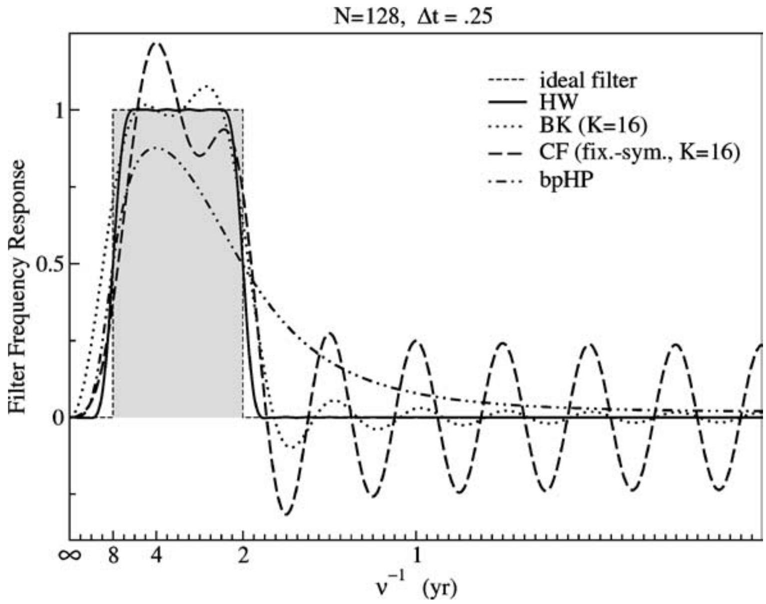


*Figure 4*. Comparison of different filters.

A nonrecursive implementation of the HP filter is however possible if we filter in the frequency domain by applying (12) to the Fourier transform of the detrended input series. In this way no spurious phase shift would be introduced. In the rest of the paper, we use the original, time-domain, recursive implementation of the HP filter.

Although the paper was published only in 1997, the HP filter dates back to 1980, and it was a tremendous innovation for that time. To our knowledge, it is still the most widely used filter in macroeconometric analysis.

### 4.2. THE BAXTER–KING FILTER

Frequency selective filters are known and used in economics since the early 1960s (Nerlove, 1964; Granger, 1964), but they experienced a great revival in the mid 1990s, thanks to the work of Baxter and King (1999) among others. The bandpass filter they propose relies on the use of a symmetric finite odd-order $M = 2K + 1$ moving average so that

$$\begin{aligned} v_j &= \sum_{n=-K}^{K} h_n\, u_{j-n} \\ &= h_0\, u_j + \sum_{n=1}^{K} h_n\, (u_{j-n} + u_{j+n}). \end{aligned} \tag{18}$$

The set of $M$ coefficients $\{h_j^{\mathrm{BK}}\}$ is obtained by truncating the ideal filter coefficients at $M$ under the constraint (3) of the correct amplitude at $\nu = 0$, that is, $H(0) = 0$ for bandpass and highpass filters and $H(0) = 1$ for lowpass filters. The BK filter coefficients thus have to solve the following optimization problem:

$$\begin{aligned} &\min_{\{h_n^{\mathrm{BK}}\}_{n=-K,\dots,K}} \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} \mathrm{d}\nu \left| \sum_{n=-K}^{K} \left(h_n^{\mathrm{BK}} - h_n^{\mathrm{ideal}}\right) e^{-i2\pi n\nu\Delta t} \right|^2, \\ &\text{s.t.} \qquad \sum_{n=-K}^{K} h_n^{\mathrm{BK}} = \frac{H(0)}{\Delta t}. \end{aligned} \tag{19}$$

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

$$h_j^{\mathrm{BK}} = h_j^{\mathrm{ideal}} + \frac{H(0) - \Delta t\, \sum_{n=-K}^{K} h_n^{\mathrm{ideal}}}{M\Delta t}. \tag{20}$$

The frequency response of the BK filter with $K = 16$ selecting the band [2, 8] years is reported in Figure 4, together with the responses of the other filters examined in this paper. The length of the BK filter is $M = 2K + 1$, so it does not depend on $N$. This implies that, unless we vary $M$, no "automatic" improvement of the filter performance is expected as more data become available.

Beside being optimal for the constrained problem (19), the BK filter has many desirable properties. First, since it is real and symmetric, it does not introduce phase shifts 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 satisfies (3), thus correctly eliminating the signal mean. Moreover, the bandpass and highpass filter response behave at least as $\nu^2$ for low $\nu$, which allows the removal of up to two unit roots (see (4) and (5)). Fourth, the filter is insensitive to deterministic linear trends, provided that $M < N$, so 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, but, if $K$ is chosen too small, the filter resolution $[(2K + 1)\Delta t]^{-1}$ would worsen. In Baxter and King (1999), a value of $K = 12$ for the passband [1.5, 8] years is found to be basically equivalent to higher values, such as 16 or 20, even if it is not the case, as is shown hereafter. As a consequence, the authors suggest putting $K \geq 12$ irrespective of $N$, the sampling frequency $\Delta t$ or 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 (19), which, by adding a constant to the ideal filter coefficients (see (20)), 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, $M$ 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 whose length $M\Delta t$ is smaller than the longest period they try to extract will perform very poorly at low frequencies. Thus, the value of $M$ must be such that

$$\nu_l > (M\Delta t)^{-1}.$$

This implies that an accurate selection of periods equal to or smaller than 8 years from quarterly data requires using *at least* a 32-point filter ($K = 16$): the low limit value of $K \geq 12$ suggested in Baxter and King (1999) in the case of a passband of [1.5, 8] years 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 with short-length time series.

Finally, it is worth noting that the same "right" behavior (3) at the origin is ensured with the truncated filter (8) with no additional constraint beyond the requirement of harmonic cutoff frequencies, that is, cutoff frequencies $\nu_{\{l,h\}}$ that are both chosen to be integer multiples of $T^{-1}$. Actually, it is easy to check from inspection of the frequency response in Figure 2 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.

### 4.3. THE CHRISTIANO–FITZGERALD FILTER

Christiano and Fitzgerald (2003) build a filter using two new ingredients: (i) accounting for the assumed spectral density of the original data, and (ii) dropping the stationarity and symmetry conditions 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:

$$\min_{\{h_j\}} \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} \mathrm{d}\nu \left| \sum_{\{j\}} \left(h_j - h_j^{\text{ideal}}\right) e^{-i2\pi j\nu\Delta t} \right|^2 |U^{\text{exact}}(\nu)|^2, \tag{21}$$

which is equivalent to the minimization of

$$\left|v_j - v_j^{\text{ideal}}\right|^2 = \int_{-(2\Delta t)^{-1}}^{(2\Delta t)^{-1}} \mathrm{d}\nu \, |H(\nu) - H^{\text{ideal}}(\nu)|^2 \, |U^{\text{exact}}(\nu)|^2 \, ; \tag{22}$$

that is, of the discrepancy between the ideally filtered data and the effectively filtered ones.

According to different types of optimization problems, which give rise to different filters, the set of indexes $\{j\}$ could be constant symmetric $j = -K, \ldots, K$, constant asymmetric $j = -K, \ldots, K'$, or even general time-varying like $j = -(N-j), \ldots, j-1$. 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 Christiano and Fitzgerald (2003)), the solution is simply given by truncating the ideal filter coefficients (7). If $u_j$ has one unit root and $|U^{\text{exact}}(\nu)|^2$ goes as $\nu^{-2}$ for low frequencies but tends to a constant at high frequencies (the *near-IID case* in Christiano and Fitzgerald (2003)), 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 (3) is satisfied. The filtering operation is

$$\begin{bmatrix} v_1^{CF} \\ v_2^{CF} \\ \\ \vdots \\ \\ v_{N-1}^{CF} \\ v_N^{CF} \end{bmatrix} = \begin{bmatrix} \hat{h} & h_1 & h_2 & \dots & h_{N-3} & h_{N-2} & h_{N-1} \\ \hat{h} - h_0 & h_0 & h_1 & \dots & h_{N-4} & h_{N-3} & \hat{h} - h_{\{0,N-3\}} \\ \hat{h} - h_{\{0,1\}} & h_1 & h_0 & \dots & h_{N-5} & h_{N-4} & \hat{h} - h_{\{0,N-4\}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hat{h} - h_{\{0,N-4\}} & h_{N-4} & h_{N-5} & \dots & h_0 & h_1 & \hat{h} - h_{\{0,1\}} \\ \hat{h} - h_{\{0,N-3\}} & h_{N-3} & h_{N-4} & \dots & h_1 & h_0 & \hat{h} - h_0 \\ h_{N-1} & h_{N-2} & h_{N-3} & \dots & h_2 & h_1 & \hat{h} \end{bmatrix} \times \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix}, \tag{23}$$

where the $h_j$'s are the ideal filter coefficients (8), $\hat{h} = h_0/2$, and we define $h_{\{0,j\}} = h_0 + h_1 + \cdots + h_j$ to simplify the notation. Remark that, the series $u_j$ in (23) has been *previously detrended by linear regression*. It is evident – as the authors themselves stress – that this filter acts differently on each date, so that we actually have $N$ different filters, represented by the rows of the matrix in (23). The CF filter is time-varying and is thus nonstationary. Moreover, as is shown in Figure 5, at each time, the coefficients are asymmetric with respect to past and future data, considering, of course, periodic boundary conditions $u_{N+j} = u_{j \bmod N}$. The asymmetry causes the CF filter response to be complex, and have a nonzero phase. Nonstationarity causes the phase, as well as the real part of the frequency response, to depend on both frequency *and* time.

The gain $G(\nu, t) = \sqrt{(\Re H_{CF}(\nu, t))^2 + (\Im H_{CF}(\nu, t))^2}$ of the CF filter frequency response and its standardized phase

$$\tilde{\Phi}_{CF}(\nu, t) = \frac{\Phi_{CF}(\nu, t)}{2\pi\nu} = \frac{1}{2\pi\nu} \arctan\left(\frac{\Im H_{CF}(\nu, t)}{\Re H_{CF}(\nu, t)}\right), \tag{24}$$

which measures the number of lags in units of $\Delta t$, are reported in Figures 6 and 7, respectively. The figures have been obtained for 128 points of quarterly data and a passband of [2, 8] years. By looking at the gain (Figure 6), we notice some leakage in the upper part of the spectrum that is particularly pronounced for at least the first and last 3 years of data, which should therefore be discarded.

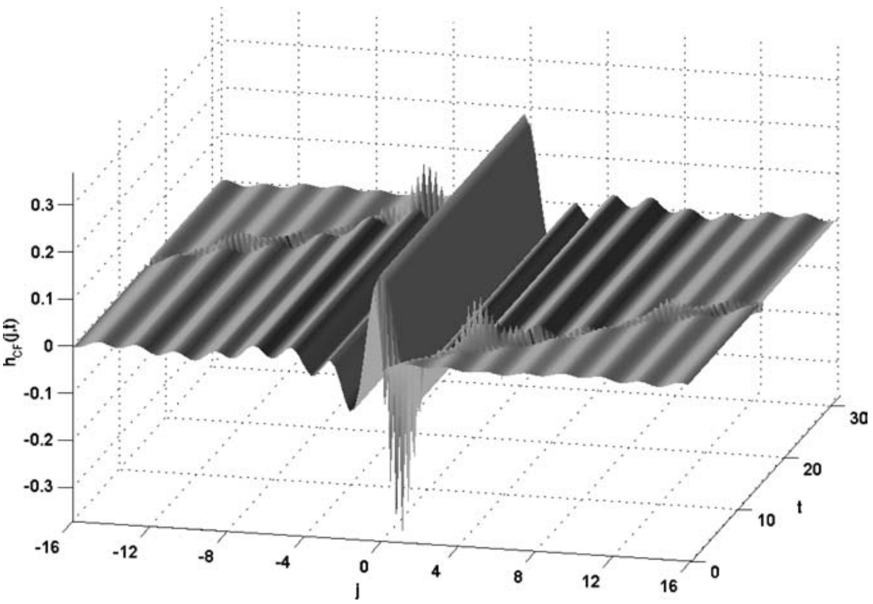


*Figure 5*. CF Random Walk Filter (I): coefficients.

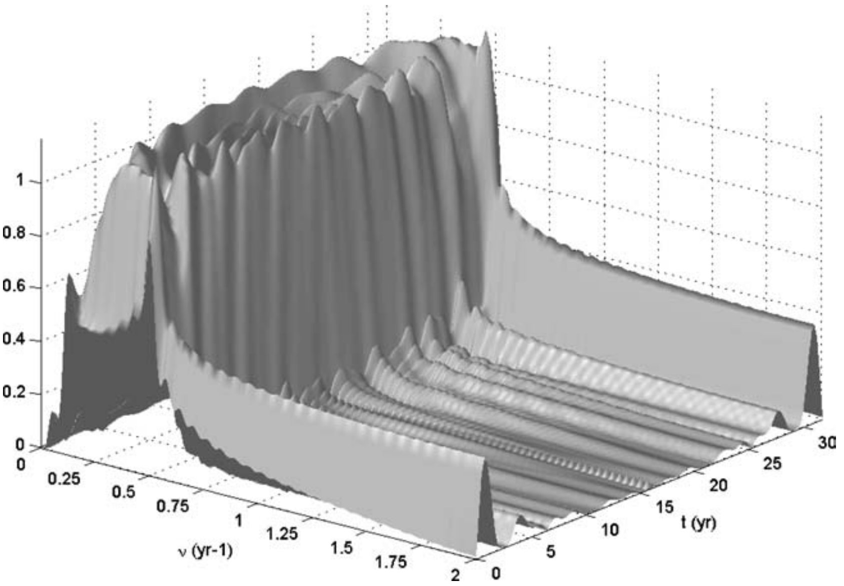


*Figure 6*. CF Random Walk Filter (II): gain.

As for the phase, Figure 7 shows the spurious shifts induced in the signal by the CF filter. The absolute value of the phase reaches a maximum of approximatively 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

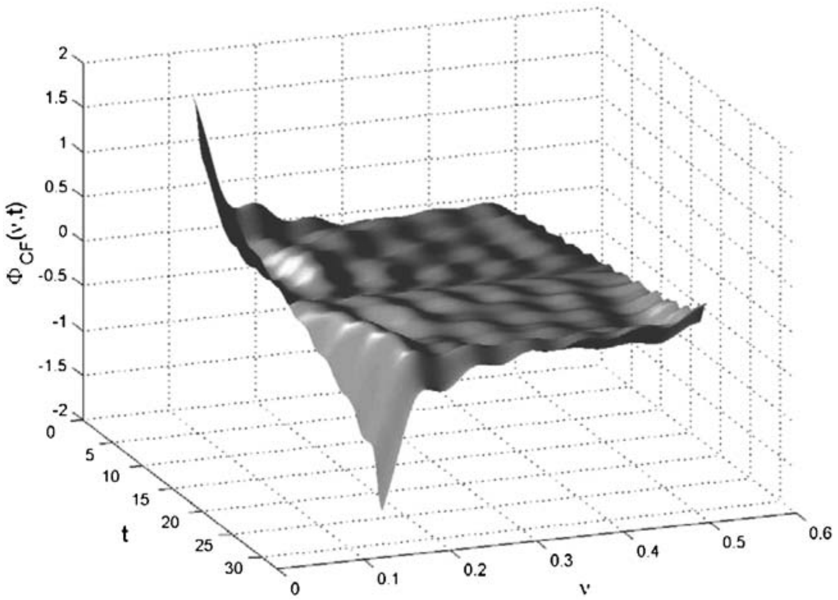


*Figure 7.* CF Random Walk Filter (IV): phase.

1 year. With the introduction of a spurious time- and frequency-dependent phase, timing and correlation properties among different frequency components within the series are irreversibly modified and cannot be recovered (see Section 5.2). The same happens to the correlation and causality relations among different series (see Section 5.3).

To make direct comparisons with the "two-dimensional" filters, we also plot the frequency response of the *symmetric fixed-length* version of the Christiano–Fitzgerald random-walk filter in Figure 4. Note that also the frequency response plots in Christiano and Fitzgerald (2003) are obtained for this version of the filter. Because of the divergence of the random-walk spectral density at low frequencies, the optimization puts strong emphasis on the behavior near the origin and attenuates these components well (see Figures 4 and 6). This comes at the price of a mediocre 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 clear whether the data can really be modeled as a random walk, as in the case of a growth rate, this filter is a rather poor approximation to the ideal one.

In our opinion, the advantages of the CF random-walk filter are minimal compared to its shortcomings. First of all, the introduction of the assumed spectral density in (21) to find the *optimal* filter adds a step in the computation of the filtered series (and not an easy one), specifically, 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 is its least problem: there is a whole stream of research on model-based filtering (Gomez, 2001; Harvey and Trimbur, 2003a). In fact, the filter finally proposed by Christiano and Fitzgerald is not optimal, but *nearly optimal*: the one they obtain by keeping the hypothesis of a random-walk for any time series. This procedure is dubious for at least two reasons: (i) the plausibility of the hypothesis, 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 Christiano and Fitzgerald (2003), p. 436, emphasis added); and (ii) if all the beauty and improvement of the method were in the search for the optimal filter for a given series, why spoil it by 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 is 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, for example, the case of the Phillips curve. A fair amount of work has been done on the issue of its "reappearance" in business cycle components (Haldane and Quah, 1999; Gaffard and Iacobucci, 2001; Stock and Watson, 1999; Christiano and Fitzgerald, 2003), but one must be aware 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 any further investigation meaningless.

Finally, we remark that nonstationary filters do not preserve purely harmonic signals, as we show in Section 5.2.

## 5. Applications

### 5.1. SPECTRAL LEAKAGE AND COMPRESSION EFFECTS

In this section, we show the effect of different filters on real data. We choose as an example the quarterly series of the Euro zone gross domestic product (Figure 8, left panel), from 1970:I to 2001:IV, thus 32 years (128 data points). The chosen band is [2, 8] years, thus the frequency responses of the applied filters are exactly those plotted in Figures 4, 6, and 7. We emphasize that we choose the abovementioned band since it is virtually equivalent to the "definition"of the business

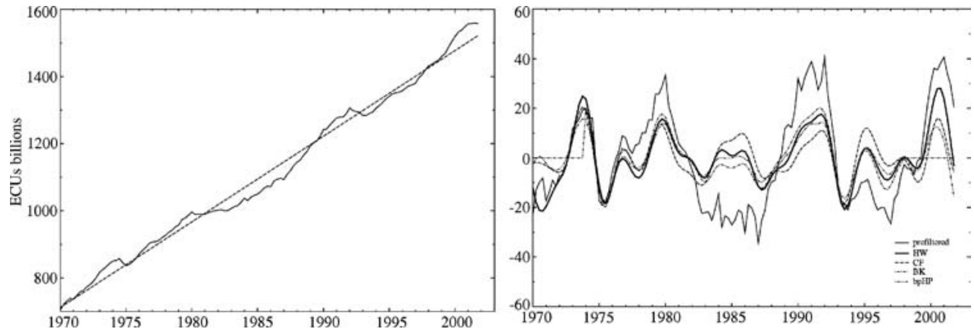


*Figure 8.* Raw and filtered Euro Zone GDP series.

cycle ([1.5, 8] years) adopted in Baxter and King (1999) following the findings in Burns and Mitchell (1946), 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 is also modified for this purpose. The condition of harmonicity of the cutoff frequencies with respect to the signal duration ensures the best performance of frequency filters. We choose the CF random-walk filter in its recommended nonsymmetric, nonstationary version. For the BK filter we choose $K = 16$ instead of 12, the value suggested in Baxter and King (1999), for reasons stated in Section 4.2. As for the HP filter, we choose its original, time-domain recursive version (11). We apply it twice, once with $\lambda \approx 677$ (677.1298) to obtain an 8 year-cutoff highpass filter and the other with $\lambda \approx 3$ (2.9142) for the 2 year-cutoff highpass filter. We subtract the second series from the first, so that we finally select the desired frequency band from the original series.

For both the HW and the CF filters, the data are linearly detrended prior to filtering. The least squares line is shown in Figure 8 (left panel); it is highly significant. The detrended series is then *prefiltered* by means of a highpass HW filter with $\nu_L = (16\,\text{years})^{-1}$ to ensure covariance stationarity and allowing the computation of a stationary spectrum (Figure 9, dotted line). This spectrum is used to define the benchmark spectrum that would issue from the application of the ideal filter on the prefiltered series. Finally, we apply the four bandpass filters to the resulting series.

The filtered series are shown in Figure 8 (right panel). Despite the fact that they have similar shapes, the amplitude of their fluctuations differs, especially near the edges. In the HP case, this is due both to the strong filter compression and to the introduced phase shift, whereas in the BK case, it is evidence of both the compression and the truncation of the coefficients. Instead, it reflects the nonstationarity of the CF filter, whose response amplitude decreases close to the ends of the data set (see Figure 6). Although harder to detect in this case than in the simple artificial series case we present in Section 5.2, the CF filter induced phase shift can be seen particularly in the periods 1980–1982 (a maximum of 1 year lead), 1986–1988

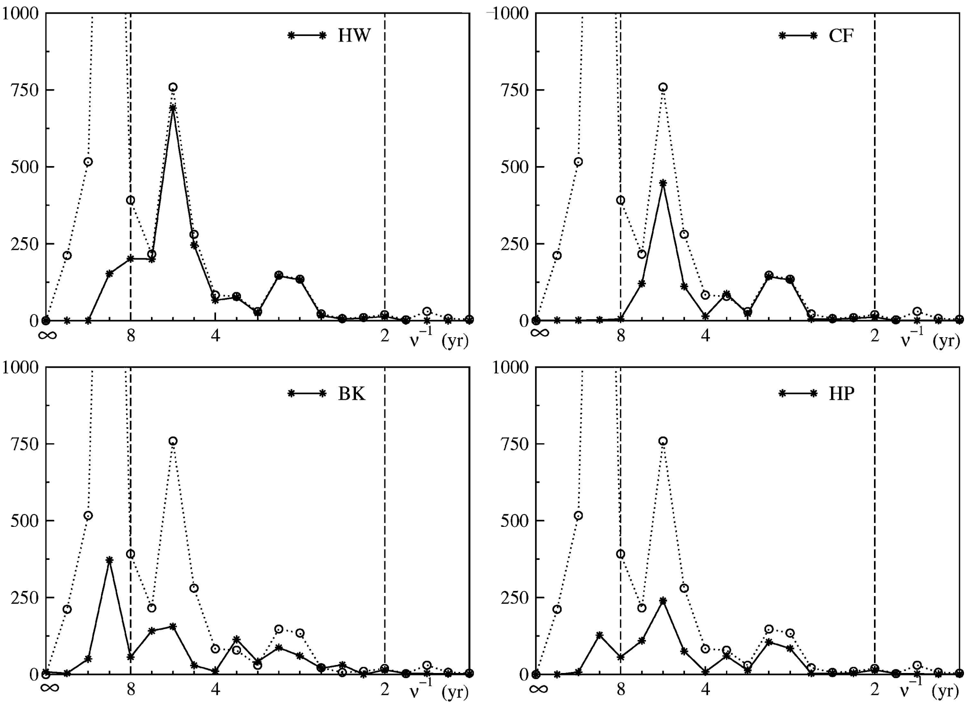


*Figure 9.* Spectra obtained by the different filtering operations.

(about 1.5 quarters lag), 1989–1991 (about 2 quarters lag) and 1995–1997 (about 3 quarters lag). Thus, in this case, the phase shift induces a stretching of the series around the point corresponding to 1984.

To assess the quality of the filtering procedures, we plot, as an estimate of the true spectra, the periodograms of each filtered series

$$P_u(k) = \Delta t \sum_{J=-(N-1)}^{N-1} \gamma_{uu}(J)\, e^{-i2\pi Jk/N}$$
$$= \Delta t \sum_{J=-(N-1)}^{N-1} \gamma_{uu}(J) \cos \frac{2\pi Jk}{N}, \tag{25}$$

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, one may object to the non-consistency of the chosen spectral estimator. Nevertheless, each filter actually operates on the periodogram, and, 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 pre-filtered in Figure 9. Our benchmark spectrum (not plotted) is null everywhere except in the passband, where it exactly matches the prefiltered series spectrum. 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 of the rectangular window. The CF filter causes significant compression in the left part of the band (top right panel). Moreover, though it performs a better attenuation than the HW filter near the lower bound of the frequency band, there is some leakage left at low frequencies (long periods), between $\nu = 0$ and $\nu = \nu_l$. 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 large leakage in the low-frequency stopband. Contrary to the previous one, the HP filter leaves the proportion among components unaffected inside the passband (bottom right panel). Furthermore, it has a much smaller leakage at low frequencies compared to the BK filter. This is visible also in Figure 4. Contrary to what the frequency responses plot shows (Figure 4), the HP filter in its recursive bandpass version seems to behave better than the BK filter. However we recall that the HP frequency response in Figure 4 corresponds to the nonrecursive symmetric implementation of the HP filter and we remind that the recursive version of the HP filter, being nonsymmetric, introduces a spurious phase-shift, as we are about to show.

### 5.2. PHASE-SHIFT EFFECTS

To test the performance of all the previously considered filters on known grounds, we apply them to an artificial series given by a simple harmonic, thus stationary, signal containing only two periods, 6 years and 1.5 years

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

where $j = 1, \ldots, 120$ and the data are supposed quarterly (30 years). We filter (26) on the band [1.5, 6] years, so that both cutoff frequencies are submultiples of the signal duration.

This procedure should be seen as a test of how the various filters behave with components that could be considered as business cycle in the Burns and Mitchell definition (Burns and Mitchell, 1946). This two-frequency series allows to highlight the filters action on each component. In fact, since these filters are not model-based, they act on any frequency independently on the structure of the input series spectrum. Furthermore, all the considered filters being *linear* operations on the data, their effect on every single frequency depends exclusively on the form of their response, that is on the number of points $N$ and the cutoff frequencies. It does not depend on the presence (or absence) of other frequency components, neither in the band nor outside it. Summing up, the presence of other frequencies in the test series would pointlessly complicate the interpretation of the filtering results and only make it harder to detect the effects we are willing to show.

The results are shown in Figure 10. The first thing we notice is that the HW-filtered series suffers less from compression than the others, while the HP-filtered series loses approximatively 50% of the original amplitude. Second, the BK and HW filters preserve the form of the input series, affecting only the amplitudes. In fact, being symmetric, both HW and BK filters perfectly preserve phase, and induce no time nor frequency shift in the filtered data. Third, ignoring for obvious reasons the BK-filtered series, the edges of the HW-filtered series follow those of the original series remarkably well, so that the harmonicity of the series is preserved, while both the CF-filtered and the HP-filtered series depart from them because of the spurious phase-shift they introduce. Fourth, the HP-filtered series shows a slight phase shift only at the edges, whereas the CF-filtered series has a progressive phase drift that affects each frequency component differently, causing the shape of the original signal to be distorted. Thus, not surprisingly, a stationary signal containing 1.5 years and 6 years components is turned, after application of the CF filter, into a

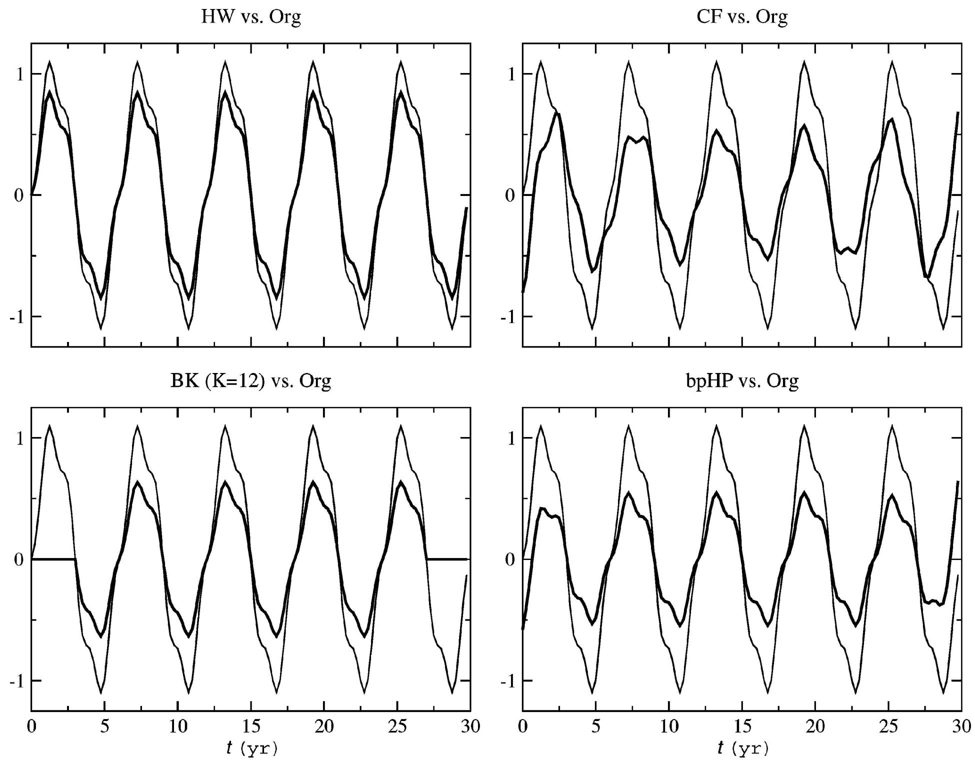


*Figure 10.* Application to an artificial series.

nonstationary one, the phase between the two components varying with time while their *mean* frequencies have been increased to $(1.49\,\text{years})^{-1}$ and $(5.7\,\text{years})^{-1}$ respectively. Moreover, the zero-points of the input series are not preserved, except for the one which is in the exact middle (15 years). Indeed, we may notice a shrink of the series around this point caused by the CF filter time-dependent phase shift.

All these shortcomings become less important when the series length is increased, but operationally the CF filter, rather than being a filter in the acknowledged sense, is best described as a smoothing procedure whose effect on frequency components can hardly be established.

### 5.3. EFFECTS ON CORRELATION

In this section, we show how the phase shift introduced by the HP and CF filters can affect the correlation properties of two series, and confront their behavior with that of the HW filter in the case of US unemployment and inflation monthly data from January '63 to December '02. We leave apart the BK filter because it induces no phase-shift, as previously verified.

Since it would be difficult to define a benchmark crosscorrelation function for the filtered series in the case of real data, we first study two artificial series, whose theoretical crosscorrelation function can be exactly computed. Once again, we simplify to the utmost the framework, by choosing two harmonic monthly series, each containing only two frequency components. The reasons of this choice are identical to those given at the beginning of Section 5.2. The two monthly series contains the periods 20 and 5 years, and 10 and 8 years respectively

$$u_{1,j} = \sin\left(\frac{2\pi j}{240}\right) - 0.15\sin\left(\frac{2\pi j}{60}\right), \tag{27}$$

$$u_{2,j} = 0.75\sin\left(\frac{2\pi j}{120}\right) + 1.15\sin\left(\frac{2\pi j}{96}\right), \quad j = 1, \ldots, 480. \tag{28}$$

Their duration is 40 years (480 points). We filter them on the band [20, 5] years using the three different filters: in this way, the theoretical crosscorrelation function between the filtered series is identical to that between the input series, since all the frequency components of both series are contained in the passband. Remark that we choose a totally different passband with respect to those chosen up to this point to show that the quality of the HW filter performance does not change when taking into account longer periodicities. This choice of the band could actually be more appropriate considering the Granger-shape spectrum of many macroeconomic series. Indeed, following Granger (1966), many researcher disagree with the Burns and Mitchell definition of the business cycle (see, among others, (Watson, 1993)), claiming that, in macroeconomics, the bulk of the series volatility is at periodicities higher than 8 years. However, this thorny issue is beyond the scope of this paper.

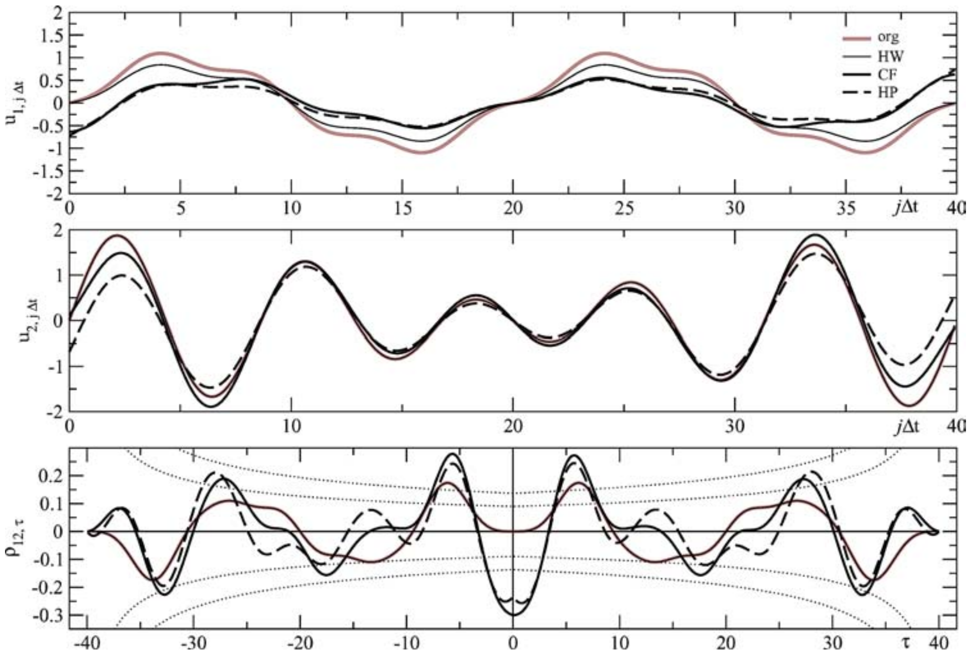


*Figure 11.* Artificial series crosscorrelation.

The results of the various filtering operations are shown in Figure 11: the upper panel shows the input series $u_{1,j}$ and the outputs of the three filtering procedures. The same for the $u_{2,j}$ series is displayed in the middle panel. In both cases the HW filter preserves the frequencies of the input series; $u_{1,j}$ is only affected in amplitude, due to the slight compression the HW filter has on the band edges, whereas $u_{2,j}$ is left completely unchanged. On the contrary, phases and frequencies are affected in both cases by the HP filter and are clearly modified in the $u_{1,j}$ case by the CF filter. This allows to anticipate that the crosscorrelations between $u_{1,j}$ and $u_{2,j}$ will be modified by the HP- and CF-filtering. In the lower panel, we plot the crosscorrelograms and their 5 and 0.5% significance levels (dotted lines). While the crosscorrelation of the HW-filtered series exactly follows that of the original one, as it should, the CF- and HP-filtered ones are wrong, differing from the correct crosscorrelation almost everywhere (and disagreeing between each other at large lags). Among the many discrepancies, the most serious is that the two series, that are uncorrelated in an interval of lags of about $\pm 2.5$ years around $\tau = 0$ – as they should, since we chose the frequencies for the series to be orthogonal at zero lag – become definitely correlated after CF or HP filtering: the maximum value of the correlation is around $-0.3$ in both cases, and it is higly significant. If we consider that the CF filter has been designed to deal with highly persistent data, which is the aim of the random walk hypoyhesis, its performance also in this case is pretty disappointing. As for the HP filter performance, we remark that its performance worsen when the chosen passband contains very low frequency components. Given this example, we are able to say that correlation properties deduced from series that have been HP- or CF-filtered should be taken with a pinch of salt.

Turning to the US Phillips curve data (Figure 12), we notice, once again, that the HW-filtered series better follows the input series at the edges, especially in the case of the unemployment rate (higher panel). In the lower panel, we plotted the crosscorrelograms of the filtered unemployment rates on the corresponding filtered inflations lagged of $\tau$ years, and the 5 and 0.5% significance levels (dotted lines). Given the results obtained in the artificial case, we take the HW-filtered series estimated correlation as the benchmark, assuming it to be closer than the others to the "true" correlation yielded by the ideal filter. The discrepancies are less evident than in the artificial example, but emerge under a careful look. Briefly mentioning some of them, both HW- and HP-filtered series have a *positive* maximum correlation value ($\tau \approx 3$ years), while the CF-filtered series have a *negative* maximum correlation value ($\tau \approx 8.33$ years). For $\tau \approx -6.7$ and $\approx 13.8$ years, the CF filter doubles the value of the correlation given by the HW filter, while at $\tau \approx 8.3$ years it gives a negative correlation larger than about one-third. Finally, at $\tau = 16.3$ years the correlation is positive for the CF-filtered series, null for the HP-, and negative for the HW-filtered series. The same thing happen around $\tau \approx -8.3$.

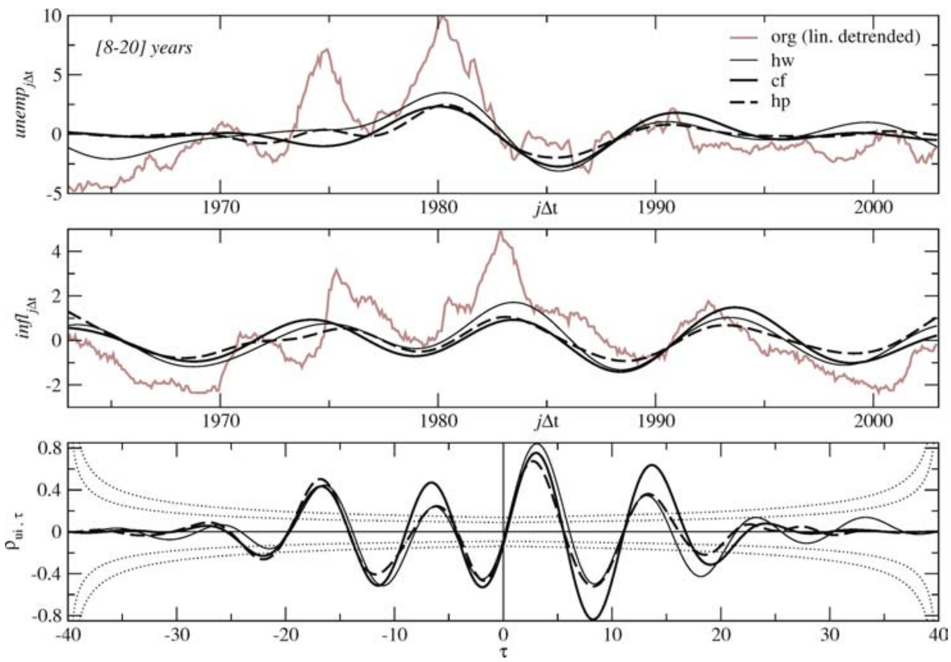


*Figure 12*. Filtered US Phillips curve data crosscorrelation.

## 6. Conclusions

To obtain an “ideal” filter, one filter selecting a finite range of frequencies with infinite resolution, requires an infinite number of data. With finite data, which is like having a lag window on an infinite-length signal, an ideal filter cannot be realized, and compromise is necessary. Simply truncating the filter coefficients to match the signal length produces a filter that 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 propose here. In particular, the HW filters 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 the time or in the frequency domain. We prefer 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 for preserving *all* timing relations among frequency components within the same series or across different series. Moreover, bandpass and highpass windowed filters are able to stationarize at least an $I(2)$ process.

The good performance of the HW filter in rejecting the off-band frequency components is checked by means of a comparison with the BK, the CF, and the HP filters. We present a critical, in-depth review of these last three filters and confront them with the HW filter. This is 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 better performing tool for the empirical study of business cycle and for establishing the correlation properties of the variables of interest. Also, we show that the correlation properties between series that have been filtered with nonstationary filters like the CF or the recursive implementation of the HP filter must be taken with caution.

Finally, for the seek of completeness, we deem it advisable to mention the main limit of these nonparametric filtering procedures, that is, their inability to extract the cycle component if the series is known to have an integrated trend (Benati, 2001; Murray, 2003). Indeed, the integrated trend spectral amplitudes are nonzero over the whole frequency range and they inevitably add to those of the cycle. Once the spectra have melted together, the only way to separate the different (trend and cycle) contributions *within* the individual frequency component is model-based filtering, since a nonparametric bandpass filter can perform separation only among *different* components. Nevertheless, if we have no prior information on the data-generating process, the nonstructural method of spectral filtering has, in our opinion, equal dignity as the model-based approach in extracting the business cycle. In the end, it is a matter of business cycle definition, a thorny subject that has a long history in empirical economics (Mills, 2003), but goes beyond the scope of this paper (Nerlove, 1964; Granger, 1964; Jenkins and Watts, 2000).

## Acknowledgement

The authors thank David Belsley, Guillaume Chevillon, Matthieu Lemoine, Stephen Pollock and Francesco Saraceno, together with two anonymous referees, for their precious comments and suggestions. 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). A `matlab` program of the filter presented in this paper is available on request.

## References

Baxter, M. and King, R.G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics*, **81**(4), 575–593.

Benati, L. (2001). *Band-pass Filtering, Cointegration and Business Cycle Analysis*. Working Paper 142, Bank of England.

Burns, A.M. and Mitchell, W.C. (1946). *Measuring Business Cycles*. Technical Report, National Bureau of Economic Research, New York.

Christiano, L. and Fitzgerald, T.J. (2003). The band-pass filter. *International Economic Review*, **44**(2), 435–465.

Gaffard, J.-L. and Iacobucci, A. (2001). The Phillips curve: Old theories and new statistics. In *Industrial Dynamics and Employment in Europe*, Scientific Report No. SOE1-CT97-1055. European Commission.

Gomez, V. (2001). The use of Butterworth filters for trend and cycle estimation in economic time series. *Journal of Business & Economic Statistics*, **19**(3), 365–373.

Granger, C.W.J. (1964). *Spectral Analysis of Economic Time Series*. Princeton University Press, Princeton, NJ.

Granger, C.W.J. (1966). The typical spectral shape of an economic variable. *Econometrica*, **34**(1), 150–161.

Haldane, A. and Quah, D. (1999). UK Phillips curves and monetary policy. *Journal of Monetary Economics*, **44**(2), 259–278.

Hamming, R.W. (1998). *Digital Filters*, 3rd edn. Dover Publications, Inc., New York.

Harvey, A.C. and Trimbur, T. (2003a). General model-based filters for extracting cycles and trends in economic time series. *The Review of Economics and Statistics*, **85**(2), 244–255.

Harvey, A.C. and Trimbur, T. (2003b). Trend estimation, signal-noise ratios and the frequency of observations. In *Proceedings of the 4th Colloquium on Modern Tools for Business Cycle Analysis, EUROSTAT*, Luxembourg.

Hodrick, R.J. and Prescott, E.C. (1997). Postwar US business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, **29**(1), 1–16.

Jenkins, G.M. and Watts, D.G. (2000). *Spectral Analysis and Its Applications*, 5th edition. Emerson-Adams Press, Inc.

King, R.G. and Rebelo, S.T. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, **17**(1–2), 207–231.

Mills, T.C. (2003). *Modelling Trends and Cycles in Economic Time Series*. Palgrave Macmillan, New York.

Murray, C.J. (2003). Cyclical properties of Baxter–King filtered time series. *The Review of Economics and Statistics*, **85**(2), 472–473.

Nerlove, M. (1964). Spectral analysis of seasonal adjustment procedures. *Econometrica*, **32**(3), 241–286.

Oppenheim, A.V. and Schafer, R.V. (1999). *Discrete-Time Signal Processing*, 2nd edn. Prentice-Hall, New Jersey.

Priestley, M.B. (1981). *Spectral Analysis and Time Series*. Academic Press, Orlando, Florida.

Ravn, M.O. and Uhlig, H. (2002). On adjusting the HP-filter for the frequency of observations. *The Review of Economic and Statistics*, **84**(2), 371–380.

Stock, J.H. and Watson, M.W. (1999). Business cycle fluctuations in US macroeconomic time series. In J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1a, chapter 1. North-Holland, Amsterdam.

Watson, M.W. (1993). Measures of fit. *Journal of Political Economics*, **101**(6), 1011–1041.