Band-Pass Filters

Martin Everts

University of Bern

January 2006

Band-Pass Filters

Martin P. Everts*

January 2006

Abstract

In the following article the ideal band-pass filter is derived and explained in order to subsequently analyze the approximations by Baxter and King (1999) and Christiano and Fitzgerald (2003). It can be shown that the filters by Baxter and King and Christiano and Fitzgerald primarily differ in two assumptions, namely in the assumption about the spectral density of the analyzed variables as well as in the assumption about the symmetry of the weights of the band-pass filter. In the article at hand it is shown that the different assumptions lead to characteristics for the two filters which distinguish in three points: in the accuracy of the approximation with respect to the length of the cycles considered, in the amount of calculable data points towards the ends of the data series, as well as in the removal of the trend of the original time series.

*University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland, Tel: +41 (0)31 631 80 76, E-Mail: martin.everts@vwi.unibe.ch. I would like to thank Ernst Baltensperger, Andreas Fischer, Roland Holder, Simon Lörtscher and Frank Schmid for helpful comments and support.

1 Introduction

If the gross domestic product of a country, a sector or an industry is analyzed "with the naked eye", clear business cycles are barely recognizable; visible is at best the long-term trend, and maybe for some time series also the short-term noise, but hardly ever the medium-term business cycles.

Therefore, mathematical methods have to be used, which divide a time series into a long-term trend, medium-term business cycles and short-term noise. Such methods are, out of obvious reasons, called filter methods or filter techniques, as they filter certain characteristics out of a time series.

In the literature several techniques are known to filter business cycles out of the gross domestic product. However, in the following only the band-pass filters by Baxter and King (1999) and Christiano and Fitzgerald (2003) are discussed. The band-pass filters by Baxter and King and Christiano and Fitzgerald are very similar in their design; they solely differ in the approximation of the ideal band-pass filter to a filter which can be applied in reality.

The article at hand firstly tries to theoretically derive the ideal band-pass filter as comprehensible as possible to further - secondly - analyze the approximation by Baxter and King (1999) and Christiano and Fitzgerald (2003). Moreover, it is attempted to highlight the differences between the two filters.

For this purpose section 2 theoretically develops the band-pass filters to subsequently discuss the approximation by Baxter and King (1999) in section 3 and the one by Christiano and Fitzgerald (2003) in section 4. Section 5 then compares the two filters and concludes.

2 Ideal Band-Pass Filter

In the following sections the ideal band-pass filter, which constitutes the basis for the approximations by Baxter and King (1999) in section 3 and the one by Christiano and Fitzgerald (2003) in section 4, is derived.

For the construction of the ideal band-pass filter, the criteria which have to be met during the extraction of the data, are defined according to Baxter and King (1999).

Table 1: Band-pass filter criteria

1. The application of an ideal band-pass should result in a stationary time series even when applied to trending data.
2. The filter should extract a specified range of periodicities.
3. The filter should leave the properties of the extracted component unaffected.
4. The ideal band-pass filter should not introduce phase shifts, i.e. it should not alter the timing relationships between series at any frequency.

These criteria apply to both, the ideal band-pass filter which will be developed in the following sections, and the approximations by Baxter and King (1999) and Christiano and Fitzgerald (2003).[1]

2.1 Stationary Time Series

In this section it is shown that the first criteria from table 1 is fulfilled, whereby a band-pass filter which is applied to data with a stochastic or a quadratic deterministic trend results in a stationary time series.

The band-pass filter is a symmetric linear filter. The observed time series y_t is thereby transformed into a new series x_t:

$$\begin{aligned} x_t &= \sum_{j=-\infty}^{\infty} \psi_j y_{t-j} \\ &= \psi(L) y_t \end{aligned} \tag{1}$$

whereas $\psi(L) = \sum_{-\infty}^{\infty} \psi_j L^j$ and L is a lag operator of the form $L^j y_t = y_{t-j}$.

To construct a filter which eliminates stochastic and quadratic deterministic trends, two assumptions for $\psi(L)$ are made. Firstly, the weights of the filter must add up to zero

$$\sum_{j=-\infty}^{\infty} \psi_j = 0$$

[1]The approximation according to Christiano and Fitzgerald (2003) does not assume a symmetric filter and therefore does not meet the first criteria from table 1.

and secondly, the weights of the filter must be symmetric

$$\psi_j = \psi_{-j}.$$

Appendix A and B proof that symmetric linear filters as described in equation 1 with weights that add up to zero have characteristics which remove stochastic and quadratic deterministic trends.

2.2 Extraction of Specific Periodicities

In a next step a filter is constructed which extracts a certain band of periodicities and hence meets the second criteria from table 1.

The main part of a band-pass filter is a so-called low-pass filter. Ideal low-pass filters pass frequencies of the band $-\underline{\omega} \leq \omega \leq \underline{\omega}$, or $|\omega| \leq \underline{\omega}$ respectively.[2]

For the creation of a low-pass filter the methodology of the spectral analysis is applied. Thereby it is assumed that a time series can be described as a weighted sum of strictly periodic processes. Moreover, the spectral representation theorem states that every time series within a broad class can be decomposed in different frequency components. The equation by Fourier (1822, p. 250) shows that an arbitrary function, which repeats itself after a certain period, consists of harmonic oscillations; that is sines and cosines functions with different phases and amplitudes and a well defined frequency. In other words: It is assumed that time series can be represented as a combination of an infinite amount of sines and cosines functions.

The equation by Fourier (1822) is often written as

$$f(\omega) = \sum_{j=0}^{\infty} (\alpha_j \cos(\omega j) + \beta_j \sin(\omega j)) \tag{2}$$

[2] To construct and ideal band-pass filter, two low-pass filters are being subtracted, so that frequencies of the band $\underline{\omega} \leq |\omega| \leq \overline{\omega}$ are let pass.

whereas $\omega = \frac{2\pi}{T}$. To transform equation 2 the Euler relations[3]

$$e^{\pm i\omega} = \cos(\omega) \pm i\sin(\omega) \tag{3}$$

are required, whereas $i = \sqrt{-1}$. The equation by Fourier can then be written as

$$f(\omega) = \sum_{j=0}^{\infty} \frac{1}{2}(\alpha_j - i\beta_j)e^{i\omega j} + \sum_{j=-\infty}^{0} \frac{1}{2}(\alpha_{-j} + i\beta_{-j})e^{-i\omega j}.$$

If this notation is extended with complex coefficients by defining ψ_j as $\frac{1}{2}(\alpha_j - i\beta_j)$ for $j > 0$, α_0 for $j = 0$ and $\frac{1}{2}(\alpha_{-j} + i\beta_{-j})$ for $j < 0$, the Fourier equation takes the following shape:

$$f(\omega) = \sum_{j=-\infty}^{\infty} \psi_j e^{i\omega j}.$$

Hence, the ideal low-pass filter has the form

$$\psi(L) = \sum_{j=-\infty}^{\infty} \psi_j L^j$$

with a frequency response function of

$$\psi(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \psi_j e^{i\omega j} \tag{4}$$

whereas $\psi_j = \frac{1}{2}(\alpha_j - i\beta_j)$ for $j \neq 0$ and $\psi_0 = \alpha_0$ for $j = 0$.[4]

2.3 No Influence

To fulfill the third criteria from table 1, whereby the time series must remain unaffected, the gain is set to 1 for those frequencies that shall be filtered out

[3]The Euler (1793) relations are often mistakenly cited as de Moivre (1722) theorem. However, the Euler relations are a derivation of the de Moivre theorem $(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$ and the exponential law $(e^{ix})^n = e^{inx}$.

[4]The condition of symmetry $\psi_j = \psi_{-j}$ leads to $\alpha_j - i\beta_j = \alpha_{-j} + i\beta_{-j}$. The constraint that the weights of the filter sum up to zero brings forth $\alpha_0 + 2\sum_{j=1}^{\infty}(\alpha_j - i\beta_j) = 0$.

and to 0 for all other frequencies.

The definition of the gain and phase of a linear filter in complex form is

$$\psi(e^{i\omega}) = |\psi(e^{i\omega})| e^{-i\theta(e^{i\omega})} \tag{5}$$

whereas $|\psi(e^{i\omega})|$ is denoted as the gain and $\theta(e^{i\omega})$ as the phase of the filter. The gain $|\psi(e^{i\omega})|$ of the filter indicates the size of the change in the amplitude of the cyclical components. The phase $\theta(e^{i\omega})$ refers to the degree of the displacement of the cyclical components.

If a time series shall remain unaffected, an ideal low-pass filter which lets frequencies between $-\underline{\omega} \leq \omega \leq \underline{\omega}$ pass, must have a gain of 1 for $|\omega| \leq \underline{\omega}$ and 0 for $|\omega| > \underline{\omega}$. Thus, the third criteria from table 1 signifies that

$$|\psi(e^{i\omega})| = \left\{ \begin{array}{lll} 1 & \text{for} & |\omega| \leq -\underline{\omega} \\ 0 & \text{elsewhere} & \end{array} \right\}. \tag{6}$$

2.4 No Phase Shifts

To fulfill the fourth criteria from table 1, whereby no phase shifts may occur, the phase $\theta(e^{i\omega})$ of the filter must be set to 0. Hence, from equation 5 it becomes apparent that $\psi(e^{i\omega})$ is symmetric; it is necessary that

$$\psi(e^{i\omega}) = \psi(e^{-i\omega}) = |\psi(e^{i\omega})|.$$

If this result is combined with equation 6, it can be shown that

$$\psi(e^{i\omega}) = \left\{ \begin{array}{lll} 1 & \text{for} & |\omega| \leq -\underline{\omega} \\ 0 & \text{elsewhere} & \end{array} \right\}.$$

2.5 Weights of the Low-Pass Filter

The individual weights ψ_n of the low-pass filter can now be calculated by means of the Fourier transformation. The derivation of the Fourier transformation can be found in appendix C. The transformation of equation 4 results in

$$\psi_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi(e^{i\omega})e^{-i\omega n}d\omega \tag{7}$$

whereas $\psi(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega}$.

As it was shown that $\psi(e^{i\omega}) = 1$ for $-\underline{\omega} \leq \omega \leq \underline{\omega}$, the individual weights from equation 23 can now be exactly determined for $n \neq 0$:

$$\begin{aligned}\psi_n &= \frac{1}{2\pi}\int_{-\omega}^{\omega} e^{-i\omega n}d\omega \\ &= \frac{1}{2\pi i n}\left(e^{i\underline{\omega}n} - e^{-i\underline{\omega}n}\right).\end{aligned}$$

If the Euler relations from equation 3 are again applied, it can be shown that $e^{i\omega} - e^{-i\omega} = 2i\sin(\omega)$ and hence, that for $n \neq 0$

$$\psi_n = \frac{1}{\pi n}\sin(\underline{\omega}n).$$

For $n = 0$, the following result holds:

$$\begin{aligned}\psi_0 &= \frac{1}{2\pi}\int_{-\omega}^{\omega} d\omega \\ &= \frac{\underline{\omega}}{\pi}.\end{aligned}$$

Recapitulating, it can be shown that the weights of a low-pass filter can be written as

$$\psi_n = \left\{\begin{array}{lll} \frac{1}{\pi n}\sin(\underline{\omega}n) & \text{for} & n \neq 0 \\ \frac{\underline{\omega}}{\pi} & \text{for} & n = 0 \end{array}\right\}$$

2.6 Weights of the Band-Pass Filter

As mentioned before, the band-pass filter is only a combination of two low-pass filters. To construct an ideal band-pass filter which allows frequencies of the band $\underline{\omega} \leq |\omega| \leq \overline{\omega}$, a low-pass filter with the frequency $\underline{\omega}$ is subtracted from a low-pass filter with the frequency $\overline{\omega}$.

Thus, if the frequency response functions of the two low-pass filters are characterized as $\psi^{\underline{LP}}(e^{i\omega})$ and $\psi^{\overline{LP}}(e^{i\omega})$, the band-pass filter has a frequency response function of

$$\psi^{BP}(e^{i\omega}) = \psi^{\overline{LP}}(e^{i\omega}) - \psi^{\underline{LP}}(e^{i\omega}).$$

According to equation 7, the weights of the band-pass filter then take the form

$$\begin{aligned}\psi_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\psi^{BP}(e^{i\omega})e^{-i\omega n}d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\psi^{\overline{LP}}(e^{i\omega})e^{-i\omega n}d\omega - \frac{1}{2\pi}\int_{-\pi}^{\pi}\psi^{\underline{LP}}(e^{i\omega})e^{-i\omega n}d\omega.\end{aligned}$$

It is now obvious that the weights of the band-pass filter can be written as

$$\psi_n = \left\{\begin{array}{ll}\frac{1}{\pi n}(\sin(\overline{\omega}n) - \sin(\underline{\omega}n)) & \text{for} \quad n \neq 0 \\ \frac{\overline{\omega}-\underline{\omega}}{\pi} & \text{for} \quad n = 0\end{array}\right\}$$

with $\overline{\omega} = \frac{2\pi}{T_{max}}$ and $\underline{\omega} = \frac{2\pi}{T_{min}}$, while T_{max} is the maximum and T_{min} the minimum duration of the cycles to be extracted.

Consequently, the ideal band-pass filter is a symmetric linear filter of the form

$$x_t = \sum_{n=-\infty}^{\infty}\psi_n y_{t-n} \tag{8}$$

with a frequency response function of

$$\psi(e^{i\omega}) = \sum_{n=-\infty}^{\infty}\psi_n e^{i\omega n} \tag{9}$$

and weights of the form

$$\psi_n = \left\{ \begin{array}{lll} \frac{1}{\pi n}\left(\sin(\frac{2\pi n}{T_{max}}) - \sin(\frac{2\pi n}{T_{min}})\right) & \text{for} & n \neq 0 \\ \frac{2}{T_{max}} - \frac{2}{T_{min}} & \text{for} & n = 0 \end{array} \right\}. \tag{10}$$

whereas T_{max} denotes the maximum and T_{min} the minimum duration of the business cycles.

However, as already mentioned at the beginning of this section the filter in equations 8 to 10 concerns an ideal band-pass filter. To calculate such a filter an infinite-order moving average would be necessary, which requires a data series of infinite lengths. As such data series do not exist in social sciences, the ideal band-pass filter must be approximated for shorter data sets.

On this point, namely in the approximation of the ideal band-pass filter, the filter by Baxter and King (1999) differs from the filter by Christiano and Fitzgerald (2003). Therefore, the approximation by Baxter and King is discussed in the following to subsequently evaluate the approximation by Christiano and Fitzgerald and ultimately to compare the two approximations.

3 Baxter-King Approximation

Baxter and King (1999) minimize the function

$$Q = \int_{-\pi}^{\pi} |\psi(e^{i\omega}) - \alpha(e^{i\omega})|^2 d\omega \tag{11}$$

whereas $\psi(e^{i\omega}) = \sum_{n=-\infty}^{\infty} \psi_n e^{i\omega n}$ is the frequency response of the ideal filter from equation 8 and $\alpha(e^{i\omega}) = \sum_{n=-K}^{K} \alpha_n e^{i\omega n}$ is the frequency response of the approximated filter. Hence, $\psi(e^{i\omega}) - \alpha(e^{i\omega})$ denotes the discrepancy of the frequency response between the approximation and the ideal filter at frequency ω. Thus, the loss function gives equal weight to the squared error terms of each individual frequency.

In their article, Baxter and King (1999) refer to Koopmans (1974) who shows that a remarkably general result for this class of optimization problems exists. The optimally approximated filter for a given maximum lag length K is constructed by cutting off the weights of the ideal filter ψ_n at the lag K. This result reflects the fact that every cut off term of a symmetric linear filter lies orthogonal to the included terms. Thus, the optimal approximated filter sets $\psi_n = 0$ for $n > K$ whereas the weights ψ_n are given in equation 10.

To calculate this approximation the function from equation 11 is minimized under the constraint that $\alpha(1) = \sum_{n=-K}^{K} \alpha_n = 0$.[5] The Lagrange function then takes the form $\mathscr{L} = Q - \lambda\alpha(1)$.

The derivation of $\mathscr{L}$ with respect to α_j results in[6]

$$\frac{\partial \mathscr{L}}{\partial \alpha_j} = -2\int_{-\pi}^{\pi} (\psi(e^{i\omega}) - \alpha(e^{i\omega}))(e^{i\omega j} + e^{-i\omega j})d\omega - 2\lambda \stackrel{!}{=} 0$$

as $\frac{\partial(\psi(e^{i\omega})-\alpha(e^{i\omega}))}{\partial \alpha_j} = -(e^{i\omega j} + e^{-i\omega j})$. If the equations for $\psi(e^{i\omega})$ and $\alpha(e^{i\omega})$ are inserted, one reaches the following result:

$$\begin{aligned} &2\sum_{n=-\infty}^{\infty} \psi_n \int_{-\pi}^{\pi} e^{i\omega n}(e^{i\omega j} + e^{-i\omega j})d\omega \\ &= 2\sum_{n=-K}^{K} \alpha_n \int_{-\pi}^{\pi} e^{i\omega n}(e^{i\omega j} + e^{-i\omega j})d\omega - 2\lambda. \end{aligned}$$

In appendix C it was shown that $\int_{-\pi}^{\pi} e^{i\omega(n-j)}d\omega = 0$ for $n \neq j$ and that $\int_{-\pi}^{\pi} e^{i\omega(n-j)}d\omega = 2\pi$ for $n = j$, as well as that $\int_{-\pi}^{\pi} e^{i\omega(n+j)}d\omega = 0$ for $n \neq -j$ and that $\int_{-\pi}^{\pi} e^{i\omega(n+j)}d\omega = 2\pi$ for $n = -j$. Thus, the derivation of $\mathscr{L}$ with respect to α_j can be written as

$$8\pi\psi_j = 8\pi\alpha_j - 2\lambda.$$

[5]Baxter and King (1999) omit the constraint in the main part of their article. However, they mention it in the appendix and use a standardization in their computer program which leads to the same result.

[6]If $\mathscr{L}$ is derived with respect to α_0, the same result arises.

The first order conditions then equal to

$$\alpha_j = \psi_j - \frac{\lambda}{4\pi}$$
$$\sum_{j=-K}^{K} \alpha_j = 0$$

and λ can be denoted as

$$\lambda = -4\pi \frac{\psi_0 + 2\sum_{n=1}^{K} \psi_n}{1+2K}$$

whereas $\sum_{n=-K}^{K} \psi_n = \psi_0 + 2\sum_{n=1}^{K} \psi_n$.

This equation system can now be solved in order to distinctly determine α_j. Summarizing, the approximated band-pass filter by Baxter and King (1999) can be written as

$$x_t = \sum_{j=-K}^{K} \alpha_j y_{t-j}$$

with a frequency response function of

$$\alpha(e^{i\omega}) = \sum_{j=-K}^{K} \alpha_j e^{i\omega j}$$

and the weights

$$\alpha_j = \psi_j - \Phi$$

while $\Phi = \frac{\psi_0 + 2\sum_{n=1}^{K} \psi_n}{1+2K}$ is a standardization factor which ensures that the weights of the approximated filter sum up to zero and hence that the approximated filter eliminates stochastic and quadratic deterministic trends. In equation 10 the weights ψ_j are defined as

$$\psi_j = \left\{ \begin{array}{ll} \frac{1}{\pi j}\left(\sin(\frac{2\pi j}{T_{max}}) - \sin(\frac{2\pi j}{T_{min}})\right) & \text{for} \quad j \neq 0 \\ \frac{2}{T_{max}} - \frac{2}{T_{min}} & \text{for} \quad j = 0 \end{array} \right\}$$

whereas T_{max} denotes the maximum and T_{min} the minimum duration of the business cycles.

4 Christiano-Fitzgerald Approximation

Christiano and Fitzgerald (2003) consider the case that y_t exists only for $t = 1, \ldots, T$ and minimize the function

$$Q = \int_{-\pi}^{\pi} \left| \psi(e^{i\omega}) - \beta(e^{i\omega}) \right|^2 f_y(\omega) d\omega. \tag{12}$$

While $\psi(e^{i\omega}) = \sum_{n=-\infty}^{\infty} \psi_n e^{i\omega n}$ denotes the frequency response function of the ideal band-pass filter from equation 8, $\beta(e^{i\omega}) = \sum_{n=t-T}^{t-1} \beta_n e^{i\omega n}$ is the frequency response function of the approximated filter, and $f_y(\omega)$ denotes the spectral density of y_t. This formulation of the loss function stresses that the solution to the problem depends on the characteristics of the time series of the filtered data, that is the spectral density.

It becomes apparent that the loss function by Baxter and King (1999) is a special case of the loss function by Christiano and Fitzgerald (2003). If it is assumed that y_t is independent and identically distributed, then $f_y(\omega) = 1$ and if further, it is assumed that the weights are symmetric, the loss function by Christiano and Fitzgerald equals the one by Baxter and King.

In contrast to the filter by Baxter and King (1999), the filter by Christiano and Fitzgerald (2003) does not assume that the weights β_j are symmetric. Therefore, according to appendix A and B, the Christiano and Fitzgerald filter does not eliminate trends in the data series y_t. Hence, before applying this filter, the stochastic or deterministic trends of a data series must be removed.

The main difference between the two filters consists in the fact that Christiano and Fitzgerald (2003) do not examine independent and identically distributed variables but an $ARMA(1, q)$ representation of the time series. In their article Christiano and Fitzgerald show that a random walk, that is an

$AR(1)$ representation of the form

$$y_t = y_{t-1} + \varepsilon_t$$

is most suitable. In the following, the approximation by Christiano and Fitzgerald is only calculated for this optimal case of a random walk. The spectral density of a random walk amounts to

$$f_y(\omega) = \frac{1}{(1 - e^{-i\omega})(1 - e^{i\omega})}.$$

Similar to Baxter and King (1999), Christiano and Fitzgerald (2003) solve the optimization problem under the condition that $\beta(1) = \sum_{n=t-T}^{t-1} \beta_n = 0$. This condition implies that $\hat{\beta}(e^{i\omega})$ is a finite-ordered polynomial, whereas $\hat{\beta}(e^{i\omega})$ is defined as

$$\hat{\beta}(e^{i\omega}) = \frac{\beta(e^{i\omega})}{1 - e^{i\omega}}$$

and $\hat{\beta}(e^{i\omega}) = \sum_{j=t-T}^{t-2} \beta_j e^{i\omega j}$. The connection between $\hat{\beta}_j$ and β_j can then be described as $\hat{\beta}_j = -\sum_{k=j+1}^{t-1} \beta_k$ or in matrix form as

$$\begin{bmatrix} \hat{\beta}_{t-2} \\ \hat{\beta}_{t-3} \\ \hat{\beta}_{t-4} \\ \vdots \\ \hat{\beta}_{t-T} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & -1 & 0 & \cdots & 0 & 0 \\ -1 & -1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ -1 & -1 & -1 & \cdots & -1 & 0 \end{bmatrix} \begin{bmatrix} \beta_{t-1} \\ \beta_{t-2} \\ \beta_{t-3} \\ \vdots \\ \beta_{t-T} \end{bmatrix} \tag{13}$$

whereas $\hat{\beta}$ is a $(T-1)$ vector and β a (T) vector. Hence, the matrix is of size $(T-1) \times (T)$.

The optimization problem from function 12 can be written as

$$Q = \int_{-\pi}^{\pi} \left| \hat{\psi}(e^{i\omega}) - \hat{\beta}(e^{i\omega}) \right|^2 d\omega$$

whereas $\hat{\psi}(e^{i\omega}) = \frac{\psi(e^{i\omega})}{1-e^{i\omega}}$ and $\hat{\beta}(e^{i\omega}) = \frac{\beta(e^{i\omega})}{1-e^{i\omega}}$. Similar to section 3 the derivation of Q with respect to β_j can be denoted as

$$\frac{\partial Q}{\partial \beta_j} = 2\int_{-\pi}^{\pi} \left(\hat{\psi}(e^{i\omega}) - \hat{\beta}(e^{i\omega})\right) e^{i\omega j} d\omega = 0$$

for $j = t-2, \ldots, t-T$, or respectively

$$\int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega j} d\omega = \int_{-\pi}^{\pi} \hat{\beta}(e^{i\omega}) e^{i\omega j} d\omega. \tag{14}$$

As already shown in section 3 the right hand side of equation 14 can be written as $2\pi\hat{\beta}_j$ where again, the results from appendix C that $\int_{-\pi}^{\pi} e^{i\omega j} d\omega = 0$ for $j \neq 0$ and $\int_{-\pi}^{\pi} e^{i\omega j} d\omega = 2\pi$ for $j = 0$, are utilized. Hence, equation 14 takes the form

$$\int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega j} d\omega \quad = \quad 2\pi\hat{\beta}_j. \tag{15}$$

Equation 15 together with the constraint that $\sum_{n=t-T}^{t-1} \beta_n = 0$ produces a system of equations with T equations and T unknowns, which, by means of equation 13, can be written in matrix form as

$$\begin{bmatrix} \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-2)} d\omega \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-3)} d\omega \\ \vdots \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-T)} d\omega \\ 0 \end{bmatrix} = 2\pi \begin{bmatrix} -1 & 0 & \cdots & 0 & 0 \\ -1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & 0 \\ 1 & 1 & \ldots & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_{t-1} \\ \beta_{t-2} \\ \vdots \\ \beta_{t-T-1} \\ \beta_{t-T} \end{bmatrix}$$

or solved for β_j as

$$\begin{bmatrix} \beta_{t-1} \\ \beta_{t-2} \\ \beta_{t-3} \\ \vdots \\ \beta_{t-T-1} \\ \beta_{t-T} \end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix} -\int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-2)} d\omega \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-2)} d\omega - \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-3)} d\omega \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-3)} d\omega - \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-4)} d\omega \\ \vdots \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-T-1)} d\omega - \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-T)} d\omega \\ \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(t-T)} d\omega \end{bmatrix}. \quad (16)$$

For $j = t-2, \ldots, t-T$ the integral $\int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega j} d\omega$ corresponds to

$$\begin{aligned} \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega j} d\omega &= \int_{0}^{\pi} \left(\hat{\psi}(e^{i\omega}) e^{i\omega j} + \hat{\psi}(e^{-i\omega}) e^{-i\omega j} \right) d\omega \\ &= \int_{\underline{\omega}}^{\overline{\omega}} \left(\frac{e^{i\omega j}}{1 - e^{i\omega}} + \frac{e^{-i\omega j}}{1 - e^{-i\omega}} \right) d\omega \end{aligned}$$

as $\hat{\psi}(e^{i\omega}) = \frac{\psi(e^{i\omega})}{1-e^{i\omega}}$ and $\psi(e^{i\omega}) = 1$ for $\underline{\omega} \leq |\omega| \leq \overline{\omega}$. For $j = 0$

$$\begin{aligned} \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) d\omega &= \int_{\underline{\omega}}^{\overline{\omega}} \left(\frac{1}{1 - e^{i\omega}} + \frac{1}{1 - e^{-i\omega}} \right) d\omega \\ &= \int_{-\pi}^{\pi} \psi(e^{i\omega}) d\omega \\ &= 2\pi\psi_0 \end{aligned} \quad (17)$$

applies, as $\frac{1}{1-e^{i\omega}} + \frac{1}{1-e^{-i\omega}} = 1$, $\psi(e^{i\omega}) = \sum_{n=-\infty}^{\infty} \psi_n e^{i\omega n}$, as well as $\int_{-\pi}^{\pi} e^{i\omega n} d\omega = 0$ for $n \neq 0$ and $\int_{-\pi}^{\pi} e^{i\omega n} d\omega = 2\pi$ for $n = 0$. For $j \neq 0$ it is noted that

$$\begin{aligned} &\int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega j} d\omega - \int_{-\pi}^{\pi} \hat{\psi}(e^{i\omega}) e^{i\omega(j+1)} d\omega \\ &= \int_{\underline{\omega}}^{\overline{\omega}} \left(\left(\frac{e^{-i\omega j}}{1 - e^{-i\omega}} + \frac{e^{i\omega j}}{1 - e^{i\omega}} \right) - \left(\frac{e^{-i\omega(j+1)}}{1 - e^{-i\omega}} + \frac{e^{i\omega(j+1)}}{1 - e^{i\omega}} \right) \right) d\omega \\ &= \int_{\underline{\omega}}^{\overline{\omega}} \left(e^{-i\omega j} + e^{i\omega j} \right) d\omega \\ &= \int_{-\pi}^{\pi} \psi(e^{i\omega}) e^{i\omega j} d\omega \\ &= 2\pi\psi_j \end{aligned} \quad (18)$$

as again $\int_{-\pi}^{\pi} e^{i\omega(j+n)}d\omega = 0$ for $j \neq -n$ and $\int_{-\pi}^{\pi} e^{i\omega(j+n)}d\omega = 2\pi$ for $j = -n$. By means of equations 17 and 18 the vector from equation 16 can be solved. Thus, the individual values for β_j are explicitly determined.

Recapitulating, the approximated band-pass filter according to Christiano and Fitzgerald (2003) can be written as

$$x_t = \sum_{j=t-T}^{t-1} \beta_j y_{t-j}$$

with a frequency response function of

$$\beta(e^{i\omega}) = \sum_{j=t-T}^{t-1} \beta_j e^{i\omega j}$$

and the weights

$$\beta_j = \left\{ \begin{array}{lll} \frac{1}{2}\psi_0 - \sum_{k=0}^{j-1} \psi_k & \text{for} & j = t-1 \\ \psi_j & \text{for} & j = t-2, \ldots, t-T-1 \\ \frac{1}{2}\psi_0 - \sum_{k=j+1}^{0} \psi_k & \text{for} & j = t-T \end{array} \right\}$$

for $t = 1, \ldots, T$. The weights ψ_k are defined in equation 10 as

$$\psi_j = \left\{ \begin{array}{lll} \frac{1}{\pi j}\left(\sin(\frac{2\pi j}{T_{max}}) - \sin(\frac{2\pi j}{T_{min}})\right) & \text{for} & j \neq 0 \\ \frac{2}{T_{max}} - \frac{2}{T_{min}} & \text{for} & j = 0 \end{array} \right\}.$$

whereas T_{max} denotes the maximum and T_{min} the minimum duration of the cycles.

5 Conclusions

The filters by Baxter and King (1999) and Christiano and Fitzgerald (2003) are based on the same ideal band-pass filter. Hence, the differences between the two filters are merely due to the type of approximation of the ideal band-pass filter. An approximation of the ideal filter is necessary as the ideal filter requires an infinite-order moving average which implies a data series of infinite length.

The approximations by Baxter and King (1999) and Christiano and Fitzgerald (2003) differ in two assumptions. The first assumption concerns the spectral density of the variables. Baxter and King assume that the variables are independent and identically distributed; Christiano and Fitzgerald presume the distribution of a random walk. The second assumption regards the symmetry of the weights of the filter. Baxter and King assume symmetric weights whereas Christiano and Fitzgerald omit this assumption.

These two dissimilarities in the assumptions of the filters by Baxter and King (1999) and Christiano and Fitzgerald (2003) lead to three divergent characteristics. The first assumption has an influence on the accuracy of the approximation with respect to the duration of the analyzed cycle. The second assumption affects on one hand the amount of output data towards the ends of the data series and on the other hand the removal of trends in the original data series.

With respect to the assumption regarding the spectral density, Baxter and King (1999) deduce independent and identically distributed variables, but Christiano and Fitzgerald (2003) assume a random walk. If the spectral density of a random walk is examined, it becomes apparent that a random walk puts more weight on lower frequencies; independent and identically distributed variables on the other hand weight all frequencies equally. Thus, it can be inferred that the filter by Christiano and Fitzgerald approximates the ideal band-pass filter for data sets with low frequencies (long durations) better than the filter by Baxter and King. However, this happens partly on costs of a worse performance in the area of high frequencies (short durations). In other words: The filter by Christiano and Fitzgerald produces more accurate results for long business cycles than the one by Baxter and King, while the filter by Baxter and King approximates the ideal band-pass filter for shorter business cycles with higher accuracy than the filter by Christiano and Fitzgerald.

The second dissimilarity - the amount of output data - results from the different assumptions with respect to the symmetry of the weights. As Baxter and

King (1999) assume symmetric weights, they are not able to make a statement about the characteristics of the K data points at the beginning and at the end of a data series, since they need these $2K$ data points for their approximation. Hence, in contrast to Christiano and Fitzgerald, Baxter and King loose K data points at the beginning and at the end of the data series. Thus, if the characteristics of the cycles towards the ends of the data series are of particular interest, it is advisable to employ the filter by Christiano and Fitzgerald.

The third difference, namely the unequal treatment of trends, can also be attributed to the assumption about the symmetry of the weights. As shown in the appendix A and B, the trend in the original time series is automatically removed if the weights of the moving average are symmetric. As Christiano and Fitzgerald (2003) make no assumption about the symmetry of the weights, the trend must be removed before applying their filter. If it is unclear whether a stochastic or a deterministic trend is at hand, it is suggested that the filter by Baxter and King (1999) is applied, as this filter assumes symmetric weights and hence automatically removes stochastic as well as deterministic trends.

Recapitulating, no clear conclusion can be drawn as for which filter should preferably be applied. However, it could be shown that the decision should depend on whether short or long business cycles are analyzed, whether the characteristics of the cycles at the beginning and at the end of the data series are of interest, and whether the trend of the original data series can be removed trouble-free.

A Removing Stochastic Trends

To demonstrate that a symmetric filter with weights which sum up to zero can eliminate stochastic trends, the filter must be rewritten as follows (whereas L^0 is defined as identity operator):

$$
\begin{aligned}
\psi(L) &= \sum_{j=-\infty}^{\infty} \psi_j L^j \\
&= \sum_{j=-\infty}^{\infty} \psi_j L^j - \sum_{j=-\infty}^{\infty} \psi_j \\
&= \sum_{j=-\infty}^{\infty} \psi_j \left(L^j - 1\right) \\
&= \sum_{j=1}^{\infty} \psi_{-j} \left(L^{-j} - 1\right) + \sum_{j=1}^{\infty} \psi_j \left(L^j - 1\right).
\end{aligned}
$$

Moreover, if the assumption of symmetric weights $\psi_j = \psi_{-j}$ is applied, the filter simplifies to

$$
\psi(L) = \sum_{j=1}^{\infty} \psi_j \left(L^j + L^{-j} - 2\right). \tag{19}
$$

Now it can be shown that

$$
L^j + L^{-j} - 2 = -(1 - L^j)(1 - L^{-j})
$$

and that

$$
\begin{aligned}
1 - L^j &= (1 - L)(1 + L + \cdots + L^{j-1}) \\
1 - L^{-j} &= (1 - L^{-1})(1 + L^{-1} + \cdots + L^{-j+1}).
\end{aligned}
$$

Thus, equation 19 can be denoted as

$$
\psi(L) = -(1 - L)(1 - L^{-1}) \sum_{j=1}^{\infty} \psi_j (1 + L + \cdots + L^{j-1})(1 + L^{-1} + \cdots + L^{-j+1}).
$$

Furthermore, it can be shown that

$$
\begin{aligned}
&(1 + L + \cdots + L^{j-1})(1 + L^{-1} + \cdots + L^{-j+1}) \\
&\quad = j + (j-1)(L + L^{-1}) + \cdots + 2(L^{j-2} + L^{-j+1}) + (L^{j-1} + L^{-j+1}) \\
&\quad = \sum_{h=-j+1}^{j-1} (j - |h|)L^h.
\end{aligned}
$$

Therefore the filter can be written as

$$
\begin{aligned}
\psi(L) &= -(1+L)(1+L^{-1}) \sum_{j=1}^{\infty} \psi_j \left(\sum_{h=-j+1}^{j-1} (j-|h|)L^h \right) \\
&= -(1+L)(1+L^{-1})\Psi_n(L)
\end{aligned}
$$

whereas $\Psi_n(L) = \sum_{j=1}^{\infty} \psi_j \left(\sum_{h=-j+1}^{j-1} (j-|h|)L^h \right)$.

If this symmetric filter is applied

$$
\begin{aligned}
x_t &= \psi(L)y_t \\
&= -(y_t + y_{t+1})(y_t + y_{t-1}) \sum_{j=1}^{\infty} \psi_j \left(\sum_{h=1-j}^{j-1} (j-|h|)y_{t+h} \right)
\end{aligned}
$$

arises. Hence, each symmetric filter $\psi(L)$ with weights that sum up to zero contains the factor $(1-L)(1-L^{-1})$. In other words: The filter $\psi(L)$ contains at least two differences. Thus, such a filter produces stationary stochastic processes of order 2 ($I(2)$).

B Removing Deterministic Trends

Hereafter, it is established that a symmetric linear filter with weights that sum up to zero does not only eliminate stochastic, but also quadratic deterministic trends. Thereto a quadratic trend

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2,$$

is specified and deployed to the symmetric filter from equation 1

$$\begin{aligned} x_t &= \sum_{j=-\infty}^{\infty} \psi_j y_{t-j} \\ &= \beta_0 \sum_{j=-\infty}^{\infty} \psi_j + \beta_1 \sum_{j=-\infty}^{\infty} \psi_j (t-j) + \beta_2 \sum_{j=-\infty}^{\infty} \psi_j (t-j)^2 \\ &= \beta_0 \sum_{j=-\infty}^{\infty} \psi_j - \beta_1 \sum_{j=-\infty}^{\infty} \psi_j j + \beta_2 \sum_{j=-\infty}^{\infty} \psi_j j^2 \\ &\quad + \left(\beta_1 \sum_{j=-\infty}^{\infty} \psi_j - 2\beta_2 \sum_{j=-\infty}^{\infty} \psi_j j \right) t + \left(\beta_2 \sum_{j=-\infty}^{\infty} \psi_j \right) t^2. \end{aligned}$$

As the weights of the filter must sum up to zero ($\sum_{j=-\infty}^{\infty} \psi_j = 0$) it follows that

$$x_t = \left(-\beta_1 \sum_{j=-\infty}^{\infty} \psi_j j + \beta_2 \sum_{j=-\infty}^{\infty} \psi_j j^2 \right) - \left(2\beta_2 \sum_{j=-\infty}^{\infty} \psi_j j \right) t.$$

Under the constraint that the weights of the filter are symmetric ($\psi_j = \psi_{-j}$) it can be shown that

$$\sum_{j=-\infty}^{\infty} \psi_j j = \sum_{j=1}^{\infty} (\psi_j - \psi_{-j}) j = 0$$

and hence that

$$x_t = \beta_2 \sum_{j=-\infty}^{\infty} \psi_j j^2.$$

Equation 20 shows that x_t is independent from t. Consequently, the filter reduces data series with a quadratic deterministic trend to series without influence on time.[7]

[7]It can also be shown that after the application of a symmetric filter with weights that sum up to zero, the trend specifications of the form $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ takes, the form $x_t = \beta_2 \sum_{j=-\infty}^{\infty} \psi_j j^2 + 3\beta_3 \sum_{j=-\infty}^{\infty} \psi_j j^2 t$, which is not independent of t.

C Fourier Transformation

In order to find the Fourier transformation of the filter

$$\psi(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \psi_j e^{i\omega j} \tag{20}$$

the two sides of the equation are multiplied with $e^{-in\omega}$ and integrated over the interval $[-\pi, \pi]$. Thus, equation 20 can be restated as

$$\begin{aligned}
\int_{-\pi}^{\pi} \psi(e^{i\omega}) e^{-in\omega} d\omega &= \int_{-\pi}^{\pi} \left(\sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega} \right) e^{-in\omega} d\omega \\
&= \sum_{j=-\infty}^{\infty} \psi_j \int_{-\pi}^{\pi} e^{i\omega(j-n)} d\omega. \qquad (21)
\end{aligned}$$

The integral on the right hand side of the equation 21 can be solved as follows:

$$\begin{aligned}
\int_{-\pi}^{\pi} e^{i\omega(j-n)} d\omega &= \left[\frac{1}{i(j-n)} e^{i\omega(j-n)} \right]_{-\pi}^{\pi} \\
&= \frac{1}{i(j-n)} \left(e^{i\pi(j-n)} - e^{-i\pi(j-n)} \right). \qquad (22)
\end{aligned}$$

Applying the Euler relations $e^{i\pi(j-n)} = \cos(\pi(j-n)) + i \sin(\pi(j-n))$ and $e^{-i\pi(j-n)} = \cos(\pi(j-n)) - i \sin(\pi(j-n))$ equation 22 can be written as

$$\begin{aligned}
\int_{\pi}^{\pi} e^{i\omega(j-n)} d\omega &= \tfrac{1}{i(j-n)} 2i \sin(\pi(j-n)) \\
&= 0 \quad \text{for} \quad j \neq n.
\end{aligned}$$

This follows as $\sin(\pi(j-n)) = 0$. For $j = n$ equation 22 can be solved by means of the Hôpital rule.

$$\begin{aligned}
\int_{\pi}^{\pi} e^{i\omega(j-n)} d\omega &= \lim_{j\to n} \frac{1}{i(j-n)} \left(e^{i\pi(j-n)} - e^{-i\pi(j-n)} \right) \\
&= \lim_{j\to n} \frac{1}{\frac{d}{dj} i(j-n)} \frac{d}{dj} \left(e^{i\pi(j-n)} - e^{-i\pi(j-n)} \right) \\
&= \lim_{j\to n} \frac{1}{i} \left(i\pi e^{i\pi(j-n)} + i\pi e^{-i\pi(j-n)} \right) \\
&= \lim_{j\to n} \pi \left(e^{i\pi(j-n)} + e^{-i\pi(j-n)} \right).
\end{aligned}$$

If the Euler relations are applied, it becomes apparent that

$$\begin{aligned}
\textstyle\int_{\pi}^{\pi} e^{i\omega(j-n)} d\omega &= \lim_{j\to n} 2\pi \cos(\pi(j-n)) \\
&= 2\pi \qquad\qquad \text{for} \quad j = n
\end{aligned}$$

as $\cos(\pi(j-n)) = 1$.

Hence, the integral equals to 0 for $j \neq n$ and 2π for $j = n$. Subsequently, equation 21 can be simplified to

$$\begin{aligned}
\int_{-\pi}^{\pi} \psi(e^{i\omega}) e^{-in\omega} d\omega &= \sum_{j=-\infty}^{\infty} \psi_j \int_{-\pi}^{\pi} e^{i\omega(j-n)} d\omega \\
&= 2\psi_n \pi.
\end{aligned}$$

This equation can now be solved for ψ_n so that

$$\psi_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi(e^{i\omega}) e^{-i\omega n} d\omega \tag{23}$$

whereas $\psi(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega}$.

References

Baxter, M., and R. G. King (1999): "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *The Review of Economics and Statistics*, 81(4), 575–593.

Christiano, L. J., and T. J. Fitzgerald (2003): "The Band Pass Filter," *International Economic Review*, 44(2), 435–65.

de Moivre, A. (1722): "De Fractionibus Algebraicis Radicalitate Immunibus ad Fractiones Simpliciores Reducendis, Deque Summandis Terminis Quarumdam Serierum Aequali Intervallo a Se Distantibus," *Philosophical Transactions*, 32, 162–178.

Euler, L. (1793): "Observationes generales circa series, quarum termini secundum sinus vel cosinus angulorum multiplorum progrediuntur," in *Nova Acta Academiae Scientiarum Imperialis Petropolitanae*, chap. 7, pp. 87–98. St. Petersburg Academy Publications.

Fourier, J. B. J. (1822): *Théorie analytique de la chaleur*. Firmin Didot, Père et fils, Paris.

Koopmans, L. (1974): *The Spectral Analysis of Time Series*. Academic Press, New York.