Separating Trends and Cycles

Takeshi Otsu

Abstract

This paper has investigated the phase-shift effects, the compression and the leakage effects of commonly-used filtering methods in macro-economics: the Baxter-King filter, the Christiano-Fitzgerald filter, the Hamming-windowed filter, and the Butterworth filter. We have used artificial data with various trend components and a cyclical component that presumably reproduce typical macroeconomic time series. The main findings are as follows. First, the Butterworth filter shows the best performance to extract components over the targeted band frequencies, when the time−series length is a multiple of the periods of the targeted frequencies. Second, the time−domain filtering indicates a severe compression. Third, the smaller the relative magnitude of the trend, the less accurate the trend estimates. Finally, the accuracy of the estimates depends on the type of trends to a lesser extent.

1 Introduction

Pre-filtering is a common practice for empirical analyses in economics to extract relevant information from economic data as accurately as possible. In macroeconomics, particularly, it attracts a great concern to identify trends and cyclical components in time series. For example, a trend component is used as an approximation to the potential output, the natural rate of employment, the structural budget deficit or the total factor productivity (TFP). A short list of studies in this line includes European Commission (1995), De Masi (1997), de Brouwer (1998), and Gerlach and Yiu (2004). In contrast, cyclical components may approximate the business cycle, and are used to investigate the statistical validity of the real business cycle models, as done by Kydland and Prescott (1982). Finally, seasonal effects may blur relations between economic variables. Then, seasonally-adjusted series might be relevant for economic analyses. The X-12-ARIMA, developed by the U.S. Census of Bureau, is one of the well-known methods frequently used by official agents in many countries. Findley, Monsell, Bell, Otto, and Chen (1998) provide a detailed explanation on the X-12-ARIMA seasonal adjustment. Another wellknown method, called the TRAMO-SEATS method, is explained in Maravall (2002). Ghysels and Osborn (2001) give a comprehensive survey of other methods and related statistical models.

In this paper, we compare filtering methods to find which methods would perform better to extract trends and business cycle components in time series in the sense that they do not give rise to statistical artifacts or spurious ‘stylized’ facts. We study the Butterworth filter in Gomez (2001) and Pollock (2000), the Hamming-windowed filter in Iacobucci and Noullez (2005), the Baxter-King (BK) filter in Baxter and King (1999), and the Christiano-Fitzgerald (CF) filter in Christiano and Fitzgerald (2003), because these filters satisfy the following desirable properties in practice. First, researchers can specify a certain periodic band of components to be extracted. The business-cycle analyses, for example, typically focus on a cyclical component with periods of one and a half to eight years. Then, researchers need to use a class of bandpass filters to extract the components of the relevant periodicities in the data. In general, the knowledge of the frequency components we put in analyses would facilitate theoretical interpretations of economic data.

Second, it is desirable that the filtering does not alter the timing relations among different time series or among frequency components within the same series. That is, it should not introduce any *phase shifts*. In most analyses, economic data are required to be aligned exactly with historical events. If pre-filtered data are subject to the phase shifts, we might misunderstand economic effects of historical events. The filtering methods considered here, except the CF filter, theoretically have no phase shifts. Although the CF filter could have phase shifts in theory, Christiano and Fitzgerald (2003, pp. 450-451) found, based on analyses of artificial data, that it showed little asymmetry.

Most widely used in the empirical literature is the Hodrick-Prescott (HP) filter, proposed by Hodrick and Prescott (1997). The HP filter is real and symmetric except at the endpoints of data. Thus, there is no phase shift in the middle parts of the filtered data, while the first and the last two data points suffer from phase shifts. Although the original HP filtering in Hodrick and Prescott (1997) does not allow researchers to extract specific periodicities, Gomez (2001) has showed that the HP filter is a two-sided Butterworth filter based on the sine function with order of two. Since a Butterworth filter is capable of tuning targeted frequencies, the HP filter is also able to extract a specified periodic range. This theoretical relation implies that the conventional HP filter for quarterly data preserves periodicities of less than 9.9 years. It would be fair to saythat the HP filter could extract desirable frequencies without significant phase shifts in practice.

We exclude the HP filter from our analyses, however. Iacobucci and Noullez (2005, pp. 84-85) have shown the HP filter has a wide transition band, and that exhibits substantial *leakage and compression*. That is, the filter might admit substantial components from the range of frequencies that were supposed to suppress (leakage), and lose substantial components over the range to be retained (compression). Otsu (2007) examines discrepancy between the ideal filter and several approximate filters, and has found that the HP filter shows great discrepancy. Therefore, it might mislead researchers to false empirical results. Harvey and Jaeger (1993) and Cogley and Nason (1995) have also pointed out that the HP filter could generate spurious business cycle dynamics. Therefore, we set it aside here.

We only consider the filtering methods that work as a descriptive tool. A safe descriptive tool would make it easy to check validity of theoretical models and prevent misinterpretation of economic implications. We can also use a class of statistical models to decompose time series into trends and cycles. For example, it can be done by the unobserved components autoregressive-integrated moving average (UCARIMA) models in Harvey (1989, p. 74). In general, model-based methods provide filters tailored to each observation set so that they satisfy a certain statistical criterion, but these filters would be different for each time series. As Sims (1974) and Wallis (1974) pointed out, the relation between two time series can be distorted when they are processed with different filters, possibly leading us to spurious empirical findings. In contrast, the filtering methods without estimation have possibility of giving a common filter to avoid the distortionary effects.

The main findings are summarized as follows. First, the Butterworth filter shows the best performance to extract components over the targeted band frequencies, when the time-series length is a multiple of the periods of the targeted frequencies. On the other hand, when the sample size is not a multiple of the periods, the frequency-domain filtering performs worse than the time-domain filtering. Extension of the samples by folding improves the performance only marginally. Second, the time-domain filtering indicates a severe compression, despite the targeted cycles are simple and deterministic. Third, the smaller the magnitude of the trend relative to that of the cycle, the less accurate the trend estimates. Finally, the accuracy of the estimates depends on the type of trends to a lesser extent. These findings suggest that it be safe to use the Butterworth filter with a sample size of a multiple of periods over the targeted frequencies. Keeping the sample length to a multiple of the periods may outweigh losing some samples.

The remaining part of this paper goes as follows. In section 2, we review filtering methods to be inspected. In section 3, we examine empirical validity of the filters with artificial data that are supposed to be typical in economic data. We make comparison among the estimates in the time domain and in the frequency domain. We also look at an overall discrepancy measure and correlations with the true values of components. Final discussion is given in section 4.

2 Bandpass Filters

We consider the following orthogonal decomposition ofthe observed series x_t :

$$x_t = y_t + \tilde{x}_t \tag{1}$$

where y_t is a signal whose frequencies belong to the interval $\{[-b, -a] \cup [a, b]\} \in [-\pi, \pi]$, while $\tilde{x}_t$ has the complementary frequencies. Suppose that we wish to extract the signal y_t. The Wiener-Kolmogorov theory of signal extraction, as expounded by Whittle (1983, Chapter 3 and 6), indicates y_t canbe written as:

$$y_t = B(L)x_t \tag{2}$$

$$B(L) = \sum_{j=-\infty}^{\infty} B_j L^j, \quad L^k x_t = x_{t-k} \tag{3}$$

In polar form, we have

$$B(e^{-iw}) = B(w) = \begin{cases} 1, & \text{for } \omega \in [-b, -a] \cup [a, b] \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $0 < a \leqq b \leqq \pi$. In application to seasonal adjustment, it is typical to set a to the seasonal frequencies concerned and b to π. In the business-cycle literature, the values of a and b are often set to the frequencies that correspond to 1.5 and 8 years, respectively. Theoretically, we need an infinite number of observations, x_t's, to compute y_t. In practice, the filtering methods approximate y_t by $\hat{y}_t$ with a finite filter. In this section, we briefly review the filtering methods to approximate y_t, which we use in the next section.

2.1 *Christiano-Fitzgerald Filter*

Suppose we seek an optimal linear approximation with finite sample observations. We find the filter weights to compute $\hat{y}_t$ to minimize the mean square error (MSE) criterion:

$$E[(y_t - \hat{y}_t)^2 \mid \mathbf{x}], \quad \mathbf{x} \equiv [x_1, ..., x_T] \tag{5}$$

Let $\hat{y}_t$ is a linear function of the observations:

$$\hat{y}_t = \sum_{j=-f}^{p} \hat{B}_j^{p,f} x_{t-j} \tag{6}$$

where $f = T - t$ and $p = t - 1$ and the $\hat{B}_j^{p,f}$'s are solution to the minimization problem of the eq.(5). Christiano and Fitzgerald (2003) express the minimization problem in the frequency domain as follows:

$$\min_{\hat{B}_j^{p,f}, j=-f,\dots,p} \int_{-\pi}^{\pi} | B(e^{-i\omega}) - \hat{B}^{p,f}(e^{-i\omega}) |^2 f_x(\omega) d\omega \tag{7}$$

where $f_x(\omega)$ is the spectral densityof x_t and 'i' indicates the imaginary number. Further,

$$\hat{B}^{p,f}(L) = \sum_{j=-f}^{p} \hat{B}_j^{p,f} L^j, \quad L^k x_t \equiv x_{t-k}$$

This is a finite approximation to eq.(3), truncating the filter length to the $p + f + 1$. In eq.(7), we express $\hat{B}^{p,f}(L)$ in polar form by replacing the lag operator L with $e^{-i\omega}$. The optimized criterion function becomes equivalent to that of Baxter and King (1999) when the data generating process is covariance-stationary with an identical and independent distribution. Therefore, the CF-type filters are derived in a more general setting than the BK filter. Christiano and Fitzgerald (2003) derived optimal weights under the following stochastic process of x_t:

$$x_t = x_{t-1} + \theta(L)\varepsilon_t, \quad E(\varepsilon_t^2) = 1 \tag{9}$$

where ε_t is white noise and $\theta(L)$ is a qth-ordered ploynomial:

$$\theta(L) = \theta_0 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q, \quad q \geqq 0 \tag{10}$$

Three points should be noticed. First, since this minimization problem depends on time t, the estimates of the signal y are computed with different filter weights, one for each date t. Second, each filter would have asymmetric lengths of past and future observations. Thus, the filtering weights are time-varying and asymmetric. When p is equal to f, the filter has symmetry and no phase shift. When both p and f are equal to a constant number (K), the filter has constant weights as well as symmetry. Then it is equivalent to the BK filter. In the latter case, the filtered time series loses $2K$ data points. Finally, when $q > 0$, we need to estimate the data generating process fo x_t to determine the value of q. Christiano and Fitzgerald (2003) estimated the moving average (MA) process of the first-differenced time series of the U.S. macroeconomic variables. Then, the estimated MA coefficients were used in the filtering procedure.

In their empirical investigations, they examined the effects of the time-varying weights, the asymmetry, and the assumption on the stochastic process. They compared variance ratios and correlations between the components extracted by the CF filters and the 'true' component. To evaluate the second moments of the 'true' component, they used the Riemann sum in the frequency domain, presuming the difference stationarity of the observations x_t's. They found that the time-varying weights and the asymmetry of the filter contributed to a better approximation, pointing out that the time-varying feature was relatively more important. Further, they claimed that the time-varying weights did not introduce severe nonstationarity in the filter approximation because the variance ratios did not vary much through the time. The correlation between $\hat{y}_t$ and y_t with different leads and lags symmetrically diminished as the leads and lags went far away, which might indicate that the degree of asymmetry was not great. Finally, the CF filter with q set to zero, which they called the Random Walk filter, gave a good approximation as much as the optimal filtering that explicitly used the estimates of the MA coefficients. Therefore, they claimed that we could use the Random Walk filter without inspecting the data generating process even if the random walk assumption was false. In this paper, we denote it by CF(RW) in the following sections.

When q is equal to zero in eq.(10), the solution to eq.(7) gives the weights of the Random Walk filter as follows:

$$B_0 = \frac{b-a}{\pi}, \quad B_j = \frac{sin(jb) - sin(ja)}{\pi j}, \quad j \geqq 1 \tag{11}$$

$$a = \frac{2\pi}{p_u}, \quad b = \frac{2\pi}{p_l} \tag{12}$$

where p_l and p_u areperiods of oscillation, satisfying $2 \leqq p_l < p_u < \infty$. The weights in eq.(11) are nothing but those of the ideal filter. Then, the Random Walk filter approximation with a smple size of T, is computed as

$$\begin{aligned}\hat{y}_t &= B_0 x_t + B_1 x_{t+1} + \cdots + B_{T-1-t}x_{T-1} + \tilde{B}_{T-t}x_T \\ &+ B_1 x_{t-1} + \cdots + B_{t-2}x_2 + \tilde{B}_{t-1}x_1 \qquad t = 3, 4, ..., T-2\end{aligned} \tag{13}$$

$$\hat{y}_2 = B_0 x_2 + B_1 x_3 + \cdots + B_{T-3}x_{T-1} + \tilde{B}_{T-2}x_T + B_1 x_1 \tag{14}$$

$$\hat{y}_{T-1} = B_0 x_{T-1} + B_1 x_T + B_1 x_{T-2} + \cdots + B_{T-3}x_2 + \tilde{B}_{T-2}x_1 \tag{15}$$

$$\hat{y}_1 = \frac{1}{2}B_0 x_1 + B_1 x_2 + \cdots + B_{T-2}x_{T-1} + \tilde{B}_{T-1}x_T \tag{16}$$

$$\hat{y}_T = \frac{1}{2}B_0 x_T + B_1 x_{T-1} + \cdots + B_{T-2}x_2 + \tilde{B}_{T-1}x_1 \tag{17}$$

where

$$\tilde{B}_{T-t} = -\frac{1}{2}B_0 - \sum_{j=1}^{T-t-1} B_j \tag{18}$$

by exploiting the fact that

$$B_0 + 2\sum_{k=1}^{\infty} B_k = 0, \tag{19}$$

2.2 *Baxter-King Filter*

Baxter and King (1999) proposed a real symmetric filter with a finite length, which is called the BK filter. Suppose that the filter has a finite length $(2K+1)$: K leads and K lags. Then, the weights are obtained by solving the following minimization problem:

$$\min_{\hat{B}_j^{K,K}, j=-K,\dots,K} \frac{1}{2\pi}\int_{-\pi}^{\pi} |B(e^{-i\omega}) - \hat{B}^{K,K}(e^{-i\omega})|^2 \, d\omega \tag{20}$$

where 'i' indicates the imaginary number as before. This is equivalent to the minimization problem in eq.(7) when the spectral density $(f_x(\omega))$ is constant, that is, the data generating process of the observations, x_t's, has the covariance-stationarity. The solution gives filter weights same as those in eq.(11), that is,

$$B_0 = \frac{b-a}{\pi}, \quad B_j = \frac{sin(jb) - sin(ja)}{\pi j}, \quad -K \leqq j \leqq K \tag{21}$$

The BK filter weights are obtained by the following normalization:

$$a_j = B_j - \frac{\sum_{h=-K}^{K} B_h}{2K+1} \tag{22}$$

Then, the approximation with T observations is computed as

$$\hat{y}_t = \sum_{h=-K}^{K} a_h x_{t-h}, \quad K+1 \leqq t \leqq T-K \tag{23}$$

Note the BK filter is real and symmetric. When researchers use this filter, they need to determine the number of leads/lags (K), losing K endpoints of data on both sides, $2K$ points in total. The filter's length is finite with $2K+1$. The truncation of the filters to the finite length causes discontinuity at the edges of the filters, creating leakage and compression in the frequency response. Further, the BK filter shows a frequency response of more than one-for-one over the passband or the stopband, which is called *exacerbation*. As the value of K increases, the ripples of the frequency response are getting small. Thus, the effects of leakage, compression, and exacerbation become reduced. Baxter and King (1999, p. 581) claimed that the number K should be more than and equal to 12 for quarterly data, while Iacobucci and Noullez (2005, p. 87) claimed that it should be at least 16. Finally, the BK filter is constrained so that the weights of the highpass or the bandpass filter are summed to zero. This condition, together with the fact that the filter is a symmetric finite odd-order moving average, guarantees to stationarize the second-order integrated process. Iacobucci and Noullez (2005, p. 87) argued that the constraint on the BK filter would cause extra discontinuity, worsening the leakage at high frequencies. According to the study of filtering distortion in Otsu (2007), the BK filter shows a comparatively large distortion: a large deviation from the ideal filters.

2.3 *Butterworth Filter*

Pollock (2000) proposed usingthe tangent-based Butterworth filters in the two-sided expression, which were called rational square-wave filters. The one-sided Butterworth filters are widely used in electrical engineering, and well documented in standard text books, such as Oppenheim and Schafer (1999) and Proakis and Manolakis (2007). The two-sided version guarantees phase neutrality (or no phase shift). The filters have finite coefficients, and its frequency response is maximally flat in the pass band; the first $(2n-1)$ derivatives of the frequency response are zero at zero frequency, when the filter has the n-th order. The filter could stationarize an integrated process of order up to $2n$. The order of the filter can be determined so that the edge frequencies of the passband and/or the stopband are aligned to the designated ones. Further, Gomez (2001) pointed out that the two-sided Butterworth filters could be interpreted as a class of statistical models called UCARIMA (the unobserved components autoregressive-integrated moving average) as explained in Harvey (1989, p. 74).

The lowpass filter is expressed as

$$BFT_L = \frac{(1+L)^n(1+L^{-1})^n}{(1+L)^n(1+L^{-1})^n + \lambda(1-L)^n(1+L^{-1})^n} \tag{24}$$

where $L^d x_t = x_{t-d}$, and $L^{-d} x_t = x_{t+d}$. Similarly, the highpass filter is expressed as

$$BFT_H = \frac{\lambda(1-L)^n(1-L^{-1})^n}{(1+L)^n(1+L^{-1})^n + \lambda(1-L)^n(1-L^{-1})^n} \tag{25}$$

Note, $BFT_L + BFT_H = 1$, which is the complementary condition required by Pollock (2000, p. 321). Here, λ is the so-called smoothing parameter. We observe that the Butterworth highpass filter in eq.(25) can handle non-stationary components integrated of order $2n$ or less. Let ω_c the *cutoff point* at which the gain is equal to 0.5. It is shown

$$\lambda = \{tan(\omega_c/2)\}^{-2n} \tag{26}$$

To see this, we replace the L by $e^{-i\omega}$ in eq.(24) to obtain the frequency response function in polar form as

$$\psi_L(e^{-i\omega};\lambda,n) = \frac{1}{1+\lambda(i(1-e^{-i\omega})/(1+e^{-i\omega}))^{2n}} \tag{27}$$

$$= \frac{1}{1+\lambda\{tan(\omega/2)\}^{2n}} \tag{28}$$

Here, it is easy to see that eq.(26) holds when $\psi_L(e^{-i\omega}) = 0.5$. We also observe in eq.(28) that the first $(2n-1)$ derivatives of $\psi_L(e^{-i\omega})$ are zero at $\omega = 0$; thus, this filter is maximally flat. Note that the gain is the modulus of the frequency response function, and indicates to what degree the filter passes the amplitude of a component at each frequency. The Butterworh filter considered here is symmetric and its frequency response function is non-negaitve. Therefore, the gain is equal to the frequency response function. Then, we can use eq.(28) to specify ω_c so that the gain at the edge of the pass band is close to one and that of the stop band close to zero. Let the pass band $[0, \omega_p]$, and the stop band $[\omega_s, \pi]$, where ω_p is smaller than ω_s. As in Gomez (2001, p. 372), we consider the following conditions for some small positive values of δ_1 and δ_2,

$$1 - \delta_1 < |\psi_L(e^{-i\omega}; \lambda, n)| \leqq 1 \quad \text{for} \quad \omega \in [0, \omega_p] \tag{29}$$

$$0 \leqq |\psi_L(e^{-i\omega}; \lambda, n)| < \delta_2 \quad \text{for} \quad \omega \in [\omega_s, \pi] \tag{30}$$

That is, we can control leakage and compression effects to some extent with the values of δ_1 and δ_2. These conditions can be written as follows:

$$1 + \left(\frac{tan(\omega_p/2)}{tan(\omega_c/2)}\right)^{2n} = \frac{1}{1-\delta_1} \tag{31}$$

$$1 + \left(\frac{tan(\omega_s/2)}{tan(\omega_c/2)}\right)^{2n} = \frac{1}{\delta_2} \tag{32}$$

Then, we determine the cutoff frequency (ω_c) and the filter's order (n), given ω_p, ω_s, δ_1 and δ_2. The closer to zeros both δ_1 and δ_2, the smaller the leakage and the compression effects. If n turns out not an integer, the nearest integer is selected.

The Butterworth filter could have another form. Instead of eq.(24), the lowpass

$$BFS_L = \frac{1}{1+\lambda(1-L)^n(1-L^{-1})^n} \tag{33}$$

$$BFS_H = \frac{\lambda(1-L)^n(1-L^{-1})^n}{1+\lambda(1-L)^n(1-L^{-1})^n} \tag{34}$$

where

$$\lambda = \{2sin(\omega_c/2)\}^{-2n} \tag{35}$$

These are the so-called sine-based Butterworth filters. When n is equal to 2, eq.(34) is the HP cyclical filter, as derived in King and Rebelo (1993, p. 224). Thus, as pointed out by Gomez (2001, p. 336), the sine-based Butterworth filter with order two ($n = 2$) can be viewed as the HP filter. As in the case of the tangent-based one, the cutoff point, ω_c, can be determined with the following conditions:

$$1+\left(\frac{sin(\omega_p/2)}{sin(\omega_c/2)}\right)^{2n} = \frac{1}{1-\delta_1} \tag{36}$$

$$1+\left(\frac{sin(\omega_s/2)}{sin(\omega_c/2)}\right)^{2n} = \frac{1}{\delta_2} \tag{37}$$

We observe that the Butterworth highpass filter in eq.(25) or eq.(34) can handle nonstationary components integrated of order $2n$ or less. The HP filter can stationarize the time series with unit root components up to the fourth order. Gomez (2001, p. 367) claimed that the BFT would give better approximations to ideal low-pass filters than the BFS. A simulation study in Otsu (2007) confirmed it. Therefore, we restrict attention to the tangent-based filters in the following.

Now, we apply the Butterworth filters to extraction of components over a certain band $[\omega_1, \omega_2]$, where ω_1 is smaller than ω_2. The bandpass filter is obtained as the difference between two highpass filters in eq.(25), or two lowpass filters in eq.(24) with different values of λ, as in Baxter and King (1999, p. 578). Suppose a lowpass filter has the pass band $[0, \omega_{p1}]$ and the stop band $[\omega_1, \pi]$. Here, ω_{p1} indicatesa frequency at which the cycle is one-period longer than at ω_1. This lowpass filter has the cutoff frequency of ω_{c1} and the order of n_1 determined in eq.(31) and (32). The corresponding value of λ is λ_1. Similarly, another lowpass filter has the pass band $[0, \omega_2]$ and the stop band $[\omega_{p2}, \pi]$. Here, ω_{p2} indicates a frequency at which the cycle is one-period shorter than at ω_2. The filter has the cutoff frequency of ω_{c2} and the order of n_2. Then, the value of λ is λ_2. The bandpass filter, $BFT^{bp}(\lambda_1, n_1, \lambda_2, n_2)$ can be obtained as

$$BFT^{bp}(\lambda_1, n_1, \lambda_2, n_2) = BFT_L(\lambda_2, n_2) - BFT_L(\lambda_1, n_1) \qquad (38)$$

The corresponding frequency response is expressed as

$$h(\omega; \lambda_1, n_1, \lambda_2, n_2) = \psi_L(e^{-i\omega}; \lambda_2, n_2) - \psi_L(e^{-i\omega}; \lambda_1, n_1) \qquad (39)$$

Alternatively, we sequentially apply the highpass filter with a lower cutoff frequency to a series, and then further apply the lowpass filter with a higher cutoff frequency to the filtered series. Pedersen (2001, p. 1096) reported that the sequential filtering had less distorting effects than the linear combination of the filters. The empirical results in the following sections do not change whether we use the difference method or the sequential method. Yet another method is to convert the lowpass filter to the bandpss filter with the highpass transformation, described in a standard textbook (e.g. Proakis and Manolakis, 2007, p. 733), and explicitly obtain the bandpss filter (see Gomez, 2001, p. 371). This filter, however, has only one order parameter, implicitly assuming n_1 is equal to n_2. As we will see later, the values of n_1 and n_2 are very different. Finally, Harvey and Trimbur (2003, pp. 248-249) derived the *generalized Butterworth bandpss filter* in the context of unobserved-component models, taking advantage of the Wiener-Kolmogorov formula. To compute the values of the smoothing parameter and the filter's order, we need determine the locational parameter value of the band and the bandwidth. Still, a numerical calculation is involved. Here, we stick to the difference method, because it is easy to control leakage and compression effects at a specific frequency.

It is possible to implement the Butterworth filtering either in the time domain or in the frequency domain. Following Pollock (2000), Otsu (2007) implemented it in the time domain, and found that the matrix inversion was so inaccurate when the cycle period was longer than seven that it was impossible to control leakage and compression effects with a certain precision specified by eq.(31) and eq.(32), or eq.(36) and eq.(37). Further, the filters at the endpoints of data are nonsymmetric due to the finite truncation of filters. This implies that the time-domain implementation introduces phase shifts. Therefore, we do not choose the time-domain filtering.

Alternatively, we can implement the Butterworth filtering in the frequency domain. The frequency-domain filtering, first, requires the Fourier transform of the observations. Suppose we have T observations, x_t. Let X_k the transformed series at the k frequency. Then, we have the discrete Fourier transform as follows:

$$X_k = \sum_{j=0}^{T-1} x_j \, e^{-i\frac{2\pi}{T}jk}, \quad k = 0, \ldots, m \tag{40}$$

$$m = \begin{cases} \frac{T-1}{2}, & \text{for odd } T \\ \frac{T}{2}, & \text{for even } T \end{cases} \tag{41}$$

In the frequency-domain filtering, the frequency response function gives filtering weights. Let $h(s)$ the frequency response function at a frequency s. For the bandpass filtering described above, we set $h(s)$ to $h(\lambda_1, n_1, \lambda_2, n_2)$ in eq.(39). Then, the approximation, $\hat{y}_j$, is computed via the inverse discrete Fourier transform as follows:

$$\hat{y}_j = \frac{1}{T}\left\{\sum_{k=0}^{m} h(k) \cdot X_k\, e^{i\frac{2\pi}{T}jk} + \sum_{k=1}^{T-1-m} h(k) \cdot X_{T-k} e^{i\frac{2\pi}{T}jk}\right\}, \tag{42}$$

$$j = 0, \cdots, T-1$$

In contrast to the time-domain implementation, the frequency-domain implementation does not introduce any phase shifts, as the theoretical backgrounds of the Butterworth filter dictate. As Baxter and King (1999, p. 580) pointed out, however, we need to remove stochastic trends which commonly exist in macroeconomic data prior to taking the Fourier transform. Then, we mustmake a choice of detrending methods, which we briefly discuss in the next section.

2.4 *Hamming-Windowed Filter*

Iacobucci and Noullez (2005) claimed that the Hamming-windowed filter be a good candidate for extracting frequency-defined components. The proposed filter has a flatter response over the passband than the HP filter, the BK filter, and the CF filter. It has almost no leakage and compression, and eliminates high-frequency components better than the other three filters. The filtering is implemented in the frequency domain. The procedure is implemented as follows. First, we subtract, if necessary, the least-square regression line to detrend the observation series, x_t's, to make it suitable for the Fourier transform. Second, we take the Fourier transform of x_t's, as in eq.(40). Third, we convolve the ideal response with a spectral window to find the windowed filter response in the frequency domain. Let the lag window $\lambda(s)$ with a truncation point $M < T - 1$:

$$\lambda(s) = \alpha + (1 - \alpha)\cos\left(\frac{\pi}{M}s\right), \quad s = -M, \cdots M \tag{43}$$

This is the so-called *General Tukey* window. $\lambda(s)$ is called the Tukey-Hamming window when α is equal to 0.54, and the Tukey-Hanning when α is 0.5 (Priestly, 1981, pp. 442-443). The corresponding spectral window at a frequency, θ, turns out

$$W(\theta) = \frac{(1-\alpha)}{2}D_M\left(\theta - \frac{\pi}{M}\right) + \alpha D_M(\theta) + \frac{(1-\alpha)}{2}D_M\left(\theta + \frac{\pi}{M}\right) \tag{44}$$

where the function $D_M(\cdot)$ denotes "*Dirichlet kernel*" (see Priestly, 1981, p. 437) given by

$$D_M(\theta) = \frac{1}{2\pi}\sum_{s=-M}^{M}\cos s\theta \tag{45}$$

Let the ideal response H for the targeted frequency range, $\{[-b, -a] \cup [a, b]\} \in [-\pi, \pi]$,

$$H(\theta) = \begin{cases} 1 & \text{if } a \leqq \theta \leqq b \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

where

$$\theta = \frac{\pi|k|}{M}, \quad k = -M, \cdots, M \tag{47}$$

To find a finite-duration impulse response (FIR) filter, we use the periodic convolution of the ideal response H with the spectral window $W(\theta)$ (see Oppenheim and Schafer, 1999, p. 466). That is,

$$W_H(\omega) = \int_{-\pi}^{\pi} H(\theta)W(\omega - \theta)d\theta \tag{48}$$

It turns out to be expressed as a simple weighted average of the values of the ideal filter at three frequencies as follows (see Priestly, 1981, pp. 433-442). For some frequency, θ,
we have

$$W_H(\theta) = \frac{(1-\alpha)}{2} H(\theta - \pi/M) + \alpha H(\theta) + \frac{(1-\alpha)}{2} H(\theta + \pi/M) \quad (49)$$

Let $h_H(k)$ the weights of the windowed response at the frequency θ in eq.(47). Then, we can rewrite eq.(49) in terms of k as follows:

$$h_H(k) = \frac{(1-\alpha)}{2} H(k-1) + \alpha H(k) + \frac{(1-\alpha)}{2} H(k+1) \quad (50)$$

Finally, we use $h_H(k)$ in eq.(42), instead of $h(k)$ to obtain $\hat{y}_t$, setting the truncation point M to m defined in eq.(41). Although both the Hanning-and the Hamming-windowed filters cause no phase shift, the latter attenuates amplitudes at low frequencies more effectively than the former. Therefore, Iacobucci and Noullez (2005) claimed that the Hamming-windowed filter would be appropriate for the short-length time series in business cycle analyses or macroeconomics, while the Hanning-windowed is for the long time series with high frequencies, typical in finance.

3 Simulation Study

In this section, we compare the filtering procedures using artificial data. The main purpose is to examine the degree of the phase shift if any, and the degree ofthe *leakage* or the *compression*. Before we get into our analyses, we discuss the drift-adjusting procedure prior to filtering. Since many economic time series involve linear or nonlinear trends, it would be better to detrend the raw data or adjust drifts of the series before filtering.

The most commonly used method is the removal of the linear regression line, recommended by Iacobucci and Noullez (2005). As shown by Chan, Hayya, and Ord (1977) and Nelson and Kang (1981), however, this method can produce spurious periodicity when the true trend is stochastic. Another widely-used detrending method is the first differencing, which reweights toward the higher frequencies and can distort the original periodicity, as pointed out by Baxter and King (1999), Chan, Hayya, and Ord (1977), and Pedersen (2001). In experimental exercises, we found these detrending methods deteriorated the filtering performance, unless the true trends happened to be same as those that the detrending methods presumed. Thus, we only present estimation results with the drift-adjusting method employed by Christiano and Fitzgerald (2003, p. 439). Let the raw data

$$x_t = z_t - (t-1)\hat{\mu} \tag{51}$$

where

$$\hat{\mu} = \frac{z_N - z_1}{N-1} \tag{52}$$

Note that $x_1 = x_N (= z_1)$. Since the discrete Fourier transform assumes circularity of the data, the discrepancy between the beginning and the end of the data may seriously distort the filtered series when we implement the filtering in the frequency domain with the Butterworth and the Hamming-windowed filters. Thus, the drift-adjusting procedure in eq.(51) would make the data suitable for filtering in the frequency domain.

We compute the cyclical component by filtering x_t. Note that the x_t of eq.(51) corresponds to that of eq.(1) in section 2. Then, we obtain the trend component by subtraction of the cyclical component from the original series z_t . Table 1 shows the parameter values of each method employed in the following exercises. We set the number of leads and lags for the BK filter to the three-year length as recommended by Baxter and King (1999, pp. 581-583 and pp. 590-591).

3.1 *Generating Artificial Data*

To examine the degree of the phase shift and the degree of the *compression*, we generate the artificial series given by the following equations:

$$z_t = T_t + C_t \ , \quad t = 1, ..., N \tag{53}$$

$$C_t = sin\left(\frac{2\pi t}{32}\right) - 0.15sin\left(\frac{2\pi t}{6}\right) \tag{54}$$

C_t indicates the cyclical component which supposedly mimics the business cycles defined by NBER (National Bureau of Economic Research), which ranges over periods of no less than 6 quarters to 32 quarters. As for a trend component, T_t , we consider five types of generating processes in Rotemberg (1999). We set parameter values so that the standard deviation of the trend component (σ_T) is 16 times as large as that of the cyclical component (σ_T), following Rotemberg (1999). These are supposed to replicate the economic time series with a strong trend. When we set the sample size N to 192, we use the following equations to generate artificial data.

$$\text{Type 1} \quad T_t = 0.2064t \tag{55}$$

$$\text{Type 2} \quad T_t = 0.2741t - (3.5439 \times 10^{-4})t^2 \tag{56}$$

$$\text{Type 3} \quad T_t = 51.4580 + 51.4580cos(1.025t/N + 3.5) \tag{57}$$

$$\text{Type 4} \quad T_t = 5.3060 + 0.1657t + 5.5231(sin(4.1t/N) - cos(4.1t/N)) \tag{58}$$

$$\text{Type 5} \quad T_t = 0.2130 + 4.2594(cos(10.25t/N)) \tag{59}$$

When we study other cases such as $N = 196$ or $\frac{\sigma_T}{\sigma_C} = 2$, the parameter values on the right hand side of these equations are adjusted by multiplying a constant value to set the ratio of the standard deviations to a designated value. These trends, together with the cycles in eq.(54), consist of the series z_t in eq.(53). Thus, we have five types of data, which are shown in Figure 1. The type 2 and the type 4 trends are concave curves, while the type 3 is convex. The type 5 series is cubic. All the periodgrams show that the 32-period cyclical component creates a sharp kink, and the 6-period cycle creates a blip at the frequency of 0.17.

3.2 *Phase Shift and Compression*

To inspect the degree of phase shifts and compressions, we plot the estimated cycles against the generated business cycles. If there is no phase shift, the estimated series have the peaks, the bottoms, and the kinks at the same data points as the original series. If there is no compression, the two series have the same magnitude. Figure 2 shows the plots of the estimates by the four methods when the sample size is 192. First of all, as theoretically expected, the Butterworth, the Hamming-windowed, and BK filtering methods do not show phase shifts. The CF(RW) filter does not show much phase shifts in the middle points, but does show notable shifts at both endpoints. These features are observed no matter what types of the trends we consider. Secondly, the frequency-domain filtering has less compression than the time-domain one. The Butterworth filter performs pretty well, while the Hamming-windowed shows a small compression. In contrast, the BK and the CF (RW) filters cause a significant compression. When the trends are nonlinear, the compression effects appear at the endpoints with the frequency-domain filtering. But, the nonlinearity of trends neither improves nor deteriorates the compression for the time-domain filtering in most cases. The exception is that the type-5 trend changes patterns of compression for the BK filtering; we observe more compression around the 120th data point.

We also look at correlations between the estimates and the generated business-cycle components and the following quantity measure of the discrepancy between them, which is used in Bruce and Gao (1996, p. 24).

$$\sqrt{\frac{\sum_{t=1+k_1}^{N-k_2}(C_t - \hat{y}_t)^2}{\sum_{t=1+k_1}^{N-k_2} C_t^2}} \qquad (60)$$

where C_t is the simulated series in eq.(54), and $\hat{y}_t$ is the estimated cycle with various methods explained in section 2. This quantity is close to zero when the discrepancy between the estimates and the generated components is small. Note that the BK filter does not produce estimates at the endpoints of data as much as the number of leads and lags. Since we set the number of leads/lags to 12, we set k_1 and k_2 to 12. Thus, we only use 168 samples through the 13th to the 180th sample point. The first panel in Table 2 shows that the Butterworth filter gives the smallest discrepancy, while the CF (RW) filter the largest. The correlation coefficients in the first panel of Table 3 show a superb performance of the frequency-domain filtering methods, and a poor performance of the CF (RW) filter. They are statistically significant because all the correlation coefficients have the zero p-values up to the fourth decimal point, which we omit from the table to save space.

One reason that the frequency-domain filtering does a better job is that the series length happens to be a multiple of the periods of the cyclical component: 32 × 6 = 192. This sample size is amenable to the frequency-domain filtering because it presumes a periodic circulation. It is interesting to see how the sample size affects the filtering performance. Here ,we present the case of 196 samples. In our experiments, this is one of the most unfavorable cases for the frequency-domain filtering. Figure 3 shows this case. It is obvious that the frequency-domain filtering performs worse than the case of 192 samples. The discrepancy measure of the second panel in Table 2 substantially deteriorates to 0.66 for the Butterworth and 0.73 for the Hamming-windowed fiters. In contrast, the time-domain methods give the discrepancy similar to the case that the sample size is a multiple of periodicities. This indicates that the periodicity assumption causes a significant artificial discontinuity at the edges of the series when we implement the filtering in the frequency domain.

As argued by Percival and Walden (2000, p. 140), it might be possible to reduce the variations at both ends of the sample if we make use of the so-called *reflection boundary treatment* to extend the series to be filtered. We notice that the transformation by eq.(51) makes both ends of the series have the same value, as mentioned before. To make periodization smooth at the endpoints, we extend the series antisymmetrically instead of symmetrically as done in the conventional reflecting rule. Let the extended series f_j .

$$f_j = \begin{cases} x_j & \text{if } 1 \leqq j \leqq N \\ 2x_1 - x_{2-j} & \text{if } -N+3 \leqq j \leqq 0 \end{cases} \quad (61)$$

That is, the $N-2$ values, folded antisymmetrically about $j=1$, are appended to the beginning of the series. We call this extension rule as the *folded extension treatment*, distinguishing from the conventional reflection. It is possible to append to the end of the series. Experimental computations show that, in both cases, the filters except the CF (RW) give much the same estimates, while the CF (RW) filter produces more accurate estimates in the former case than the latter. This is one reason that we adopt eq.(61). Another reason is that most filters give accurate and stable estimates over the middle range of the series. Since the initial data point indicates the farthest past in the time series, it does not make sense that the estimate of the initial point is subject to a large revision when additional observations are obtained in the future periods. Thus, we put the initial points in the middle part of the series to be filtered. The filtering results with the extended series are presented in Figure 4. We observe a moderate reduction in compression effects with the Butterworth and the Hamming-windowed filters. The third panel in Table 2 also shows a reduction of the discrepancy. To inspect the compression effects near the frequency band, we enlarge the lower-frequency edge of the pass band from $\frac{2\pi}{32}$ to $\frac{2\pi}{34}$; we include two-period longer cycles. The estimation results are graphed in Figure 5. Although the BK filter does not improve estimation quality, other three filters give rise to substantially improved estimates. This implies a substantial compression over the targeted frequency range. The CF (RW) filter provides the best estimates among others, tracking the generated cycle pretty well.

The overall discrepancy measure in the fourth panel of Table 2 also documents how a large compression is caused by filtering. Widening the designated bands improves accuracy of the estimation. The correlation coefficients of the estimates with the generated cycles are shown in the second and the third panels of Table 3 when using the 196 samples. It is interesting that, in terms of the correlation, the sample extension and the widened band do not much improve the filtering performance of the Butterworth, the Hamming-windowed, and the BK filters, while the CF (RW) filter shows a substantial improvement.

In practice, however, we cannot widen the pass band because the filtering might pick up the periodic cycles that should be suppressed, leading to a large leakage effect. Thus, it matters to keep a sample size to a multiple of the targeted periodicity. In experiments, we also examine the cyclical component composed of the 3-, 5-, 7- and 13-period cycles. If the sample size is $1365 (= 3 \times 5 \times 7 \times 13)$, the frequency-domain filtering, particularly with the Butterworth filter, gives good estimates. When the sample size deviates from a multiple of the periodicity, the estimation accuracy deteriorates substantially. This result is robust with various trend types.

3.3 *Trend Components*

In this section, we compare the estimated trend components with the generated true trends. We estimates the trend components by subtracting the cyclical components from the original series. Therefore, we expect a filtering performance similar to the case of the cycle estimates. Here, we only present selected results to save space. We choose a simple linear trend case, data type 1, and one of the nonlinear trend cases, data type 4, with which the filtering methods give the worst results in the previous subsection.

Figure 6 presents the case of the 192 sample size for data type 1. The left-hand side of the figure shows the power spectrum in the frequency domain. The estimated 32-period-cycle point is labeled. The Butterworth filter produces the same spectrum as the generated trend, except that there is a slight deviation at the 32-period-cycle point. The largest deviation is observed for the CF (RW) filter on the last panel on the left-hand side. The estimated trends on the right-hand show that the Butterworth filter gives the best estimates, the light-coloured circles (the estimates) lying atop the solid line (the original values). The estimates with the Hamming-windowed filter ripple slightly. The time-domain filtering with the BK and the CF (RW) filters generates a moderate wave. We have a similar observation in Figure 7 for data type 4. The largest wave is caused by the CF (RW) filter. The correlations between the estimates and the original values also document the same finding. The first panel of Table 4 shows that the correlation of the Butterworth estimates with the true trend is one, and that of the CF (RW) is lowest, although still high (0.9993).

With the 196 samples, the trend estimates are less accurate as expected from the previous findings on the estimates of the cyclical components. The correlations in the second panel of Table 4 are lower than the first panel for the frequency-domain filtering, and same for the time-domain filtering. When filtering with the series extended by folding, we can improve Butterworth estimates so that they become comparable with the CF (RW) estimates (the third panel), but we observe no significant changes for the Hamming-windowed and the Baxter-King filters. The related graphs are presented in Figure 8 for data type1and Figure 9 for data type 4. The spectrum of the CF (RW)-filter estimates gives the best fit to that of the original trend except at the 32-period frequency, showing the least dependence of the periodic circulations. In the time-domain plots on the right-hand side, all the estimates ripple in both figures. Particularly, the Hamming-windowed estimates apparently ripple to a greater degree. This indicates a leakage effect. As an experiment, we also exclude the lower-frequency component up to the 34-period cycles, resulting in the correlations in the fourth panel of Table 4. Then, all the filters but the BK filter give higher correlations. It implies the leakage from the cyclical component to the trend component, except the case of the Baxter-King filtering that does not show any improvement with the extended series and the expanded frequency range.

Up to this point, we work with the sample generated so that the magnitude of the trend component is 16 times as large as that of the cyclical component ($\frac{\sigma_T}{\sigma_C} = 16$). This is supposed to replicate a strong trend that is typical in many macroeconomic time series. When we are concerned with extraction of a trend component, however, the relative magnitude matters. The smaller the relative magnitude ($\frac{\sigma_T}{\sigma_C}$), the larger the observed leakage effect. Here, we present the case that $\frac{\sigma_T}{\sigma_C}$ equals to 2. The generated series for data type 1 and data type 4 are plotted in Figure 10, showing the significant nature of the cycles. Figure 11 and Figure 12 show the estimation results with 192 samples for data type 1 and for data type 4, respectively. The Butterworth filtering produces only small ripples, while other three filtering methods show large waves in both data types. When we use 196 samples, however, the frequency-domain filtering substantially deteriorates its performance, and the time-domain filtering does not produce much different estimates from those with the 192 samples. As shown in Figure 13 and Figure 14, the data extension does not improve accuracy of the estimates. When we widen the frequency range to extract the cyclical component, the amplitude of the waves shrinks to a greater extent as seen in Figure 15 and Figure 16. This implies a large leakage effect. These findings are confirmed by correlation values in Table 5: substantial improvement of the correlation values.

4 Discussion

This paper has investigated the phase-shift effects, the compression and the leakage effects of the filtering methods used in macroeconomics: the Baxter-King filter, the Christiano-Fitzgerald filter, the Hamming-windowed filter, and the Butterworth filter. We have used artificial data with various trend components and a cyclical component that presumably replicate typical macroeconomic time series. The main findings are summarized as follows. First, the Butterworth filter shows the best performance to extract components over the targeted band frequencies, when the time-series length is a multiple of the periods of the targeted frequencies. On the other hand, when the sample size is not a multiple of the periods, the frequency-domain filtering does worse performance than the time-domain filtering. Extension of the samples by folding improves the performance only marginally. Second, the time-domain filtering indicates a severe compression, despite the targeted cycles are simple and deterministic. Third, the smaller the magnitude of the trend relative to that of the cycle, the less accurate the trend estimates. Finally, the accuracy of the estimates depends on the type of trends to a lesser extent. These findings suggest that it be safe to use the Butterworth filter with a sample size of a multiple of periods with conspicuous gains over the targeted frequencies. Keeping the sample length to a multiple of the periods may outweigh losing some samples.

As long as economic models intend to explain a part of economic data or no econometric models can handle all the components without misspecifications, pre-filtering would be a useful method for economic analysis. There are, however, pros and cons about decomposition of economic data into trend and cyclical components. We briefly discuss them as a final remark. Singleton (1988) claimed that seasonal adjustment would distort the growth rates of economic time series, examining the Granger causality of macroeconomic data. It also pointed out that the pre-filtering would change the Euler equation derived by intertemporal optimization in the context of equilibrium business cycle models, even if the same filter were applied to all the variables. Further, if the filters were two-sided and agents' information set included them, the conventional orthogonal conditions might not hold. Finally, when economic shocks affect optimal economic decision making via propagation mechanisms in economic models, it is not appropriate to consider that the trend and cyclical components are separable. These considerations lead Singleton (1988) to prefiltering of the data would generate inconsistent estimates of model parameters. Further, Ghysels (1988) showed that high adjustment costs, which characterize the intertemporal dynamics of a model, should cause seasonal components to have a considerable power at the nonseasonal frequency in equilibrium. Then, the decomposition was not justified. King and Rebelo (1993, p. 225) also pointed out that the separability required an implausible condition either that the cycle consisted of uncorrelated events or that the trend and the cyclical components had an identical propagation mechanism of shocks. Therefore, such a decomposition would give erroneous interpretations of business cycles.

In contrast, Sims (1974) and Wallis (1974) showed that seasonally adjusted data could generate estimates of a distributed lag regression model with smaller approximation bias than unadjusted data. Sims (1993) verified in the context of rational expectations models that when seasonality was highly predictable, use of adjusted data would create estimation bias to a lesser extent. Further, Hansen and Sargent (1993) showed numerically that there was no loss of statistical consistency in estimating a correctly-specified model with seasonally adjusted data, and large reductions in asymptotic bias in case that the model was misspecified. Their results showed, however, that if a true model incorporated a seasonal habit persistence, ignoring such a persistence could result in a large estimation bias with seasonal adjusted data. Therefore, if propagation mechanisms critically depend on seasonality of adjustment costs, preferences or technology, we may have no improvement in estimation with seasonally adjusted data. All these arguments would be valid when we use filtering methods for the purpose other than seasonal adjustment. There remains much to be done to judge how useful the filtering methods in economics are. Hopefully, the findings here have given some guidance to safe use of the filtering methods.

Acknowledgements **It is gratefully acknowledged that this work is supported by JSPS (Japan Society for the Promotion of Science) KAKENHI(17530223). I also notice a grant from Seijo University under the title of「周期的経済変動抽出方法の比較分析」.**

References

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

Bruce, Andrew, and Hong-Ye Gao, 1996, Applied Wavelet Analysis with S-PLUS, Springer-Verlag, New York.

Chan, Hung C., Jack C. Hayya, and J. Keith Ord, 1977, “A note on trend removal methods: The case of polynomial regression versus variate differencing,” Econometrica 45, 737-744.

Christiano, Lawrence J. and Terry J. Fitzgerald, 2003, “The band pass filter,” International Economic Review 44 (2), 435-465.

Cogley, Timothy and James M. Nason, 1995, “Effects of the Hodrick-Prescott filter on trend and difference stationary time series, implications for business cycle research,” Journal of Economic Dynamics and Control 19, 253-278.

de Brouwer, Gordon, 1998, “Estimating output gaps,” Research Discussion Paper 9809, Economic Research Department, Reserve Bank of Australia, August.

De Masi, Paula R., 1997, “IMF estimates of potential output: Theory and practice,” IMF Working Paper 177.

European Commission, 1995, “Technical note: The commission services’ method for the cyclical adjustment of government budget balances,” Broad Economic Policy Guidelines, European Economy No. 60, 35-49.

Findley, David F., Brian C. Monsell, William R. Bell, Mark C. Otto, and Bor-Chung Chen, 1998, “New capabilities and methods of the X-12-ARIMA seasonal-adjustment program,” Journal of Business and Economic Statistics 16 (2), 127-155.

Gerlach, Stefan and Matthew S. Yiu, 2004, “Estimating output gaps in Asia: A cross-country study,” Journal of the Japanese and International Economies 18, 115-136.

Ghysels, Eric, 1988, “A study toward a dynamic theory of seasonality for economic time series,” Journal of the American Statistical Association 83, 168-172.

Ghysels, Eric and Denise R. Osborn, 2001, The econometric analysis of seasonal time series, Cambridge University Press, U.K.

Gomez, Victor, 2001, “The use of Butterworth filters for trend and cycle estimation in economic time series,” Journal of Business and Economic Statistics 19 (3), 365-373.

Hansen, L.P. and T.J. Sargent, 1993, “Seasonality and approxmimation errors in rational expectaions models,” Journal of Econometrics 55, 21-55.

Harvey, A.C., 1989, Forecasting, Structural Time Series Models and the Kalman filter, Cambridge University Press.

Harvey, A.C. and A. Jaeger, 1993, “Detrending, stylized facts, and the business cycle," Journal of Applied Econometrics 8, 231-247.

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

Hodrick, Robert J. and Edward C. Prescott, 1997, "Post-war U.S. business cycles: An empirical investigation," Journal of Money, Credit, and Banking 29, 1-16. Also in DiscussionPaper No. 451, Department of Economics, Carnegie-Mellon University, Pittsburgh, PA, 1980.

Iacobucci, Alessandra and Alain Noullez, 2005, "A frequency selective filter for short-length time series," Computational Economics 25, 75-102.

King, Robert G. and Serigio T. Rebelo, 1993, "Low frequency filtering and real business cycles," Journal of Economic Dynamics and Control 17, 207-231.

Kydland, Finn and Edward C. Prescott, 1982, "Time to build and aggregate fluctuations," Econometrica 50, 1345-1370.

Maravall, A., 2002, "Brief description of the TRAMO-SEATS methodology," in Modeling Seasonality and Periodicity, ed. Y. Kawasaki, Tokyo Japan, the Institute of Statistical Mathematics, pp. 23-40.

Nelson, Charles R. and Heejoon Kang, 1981, "Spurious periodicity in inappropriately detrendedtime series," Econometrica 49, 741-751.

Oppenheim, Alan V. and Ronald W. Schafer, 1999, Discrete-time signal processing 2nd edition, Prentice Hall International.

Otsu, Takeshi, 2007, "Time-invariant linear filters and Real GDP: A case of Japan," Seijo University Economic Papers 174, 29-47, The Economic Institute of Seijo University (Japan).

Pedersen, Torben Mark, 2001, "The Hodrick-Prescott filter, the Slutzky effect, and the distortinary effect of filters," Journal of Economic Dynamics and Control 25, 1081-1101.

Percival, Donald B. and Andrew T. Walden, 2000, Wavelet Methods For Time Series Analysis, Cambridge University Press, Cambridge.

Pollock, D.S.G., 2000, "Trend estimation and de-trending via rational square-wave filters," Journal of Econometrics 99, 317-334.

Priestly, M.B., 1981, Spectral Analysis and Time Series, Academic Press.

Proakis, John G. and Dimitris G. Manolakis, 2007, Digital Signal Processing 4th edition, Pearson Prentice Hall.

Rotemberg, Julio J., 1999, "A heuristic method for extracting smooth trends from economic times series," NBER Working Paper 7439.

Sims, C.A., 1974, "Seasonality in regression," Journal of the American Statistical Association 69, 618-626.

Sims, C.A., 1993, "Rational expectations modeling with seasonally adjusted data," Journal of Econometrics 55, 9-19.

Singleton, K.J., 1988, "Econometric issues in the analysis of equilibrium business cycle models," Journal of Monetary Economcis 21, 361-386.

Wallis, K.F., 1974, "Seasonal adjustment and relations between variables," Journal of the American Statistical Association 69, 18-31.

Whittle, Peter, 1983, Prediction and regulation by linear least-square methods, 2nd ed. (revised), Basil Blackwell, Oxford.

Table 1 Parameter Values for Filtering

Filtering Methods	Bandpass Parameters			Tuning Values
CF (Random Walk)	$a = \frac{2\pi}{32}$	$b = \frac{2\pi}{6}$		$q = 0$
Baxter-King	$a = \frac{2\pi}{32}$	$b = \frac{2\pi}{6}$		$K = 12$
Hamming-windowed	$a = \frac{2\pi}{32}$	$b = \frac{2\pi}{6}$		$\alpha = 0.54$
Butterworth	$\omega_{p1} = \frac{2\pi}{33}$	$\omega_1 = \frac{2\pi}{32}$, $\omega_2 = \frac{2\pi}{6}$	$\omega_{p2} = \frac{2\pi}{5}$	$\delta_1 = \delta_2 = 0.07$

Explanations of the parameters in Section 2.

Table 2　Discrepancy between Estimates and True Values

Sample Range: 13-180, Sample Size: 168

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.0715	0.2311	0.4224	0.5495
Data type 2	0.0818	0.2347	0.4229	0.5505
Data type 3	0.0819	0.2322	0.4222	0.5521
Data type 4	0.1065	0.2431	0.4256	0.5602
Data type 5	0.0702	0.2208	0.4987	0.5214

Sample Ranee: 13-184, Sample Size: 172

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.6660	0.7340	0.4224	0.5478
Data type 2	0.6662	0.7333	0.4236	0.5494
Data type 3	0.6670	0.7357	0.4220	0.5494
Data type 4	0.6679	0.7335	0.4253	0.5606
Data type 5	0.6664	0.7357	0.4806	0.5218

Sample Range: 13-184, Sample Size: 172 (folded extension)

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.4912	0.7092	0.4224	0.4812
Data type 2	0.4906	0.7087	0.4236	0.4912
Data type 3	0.4919	0.7095	0.4220	0.4717
Data type 4	0.4926	0.7098	0.4253	0.5088
Data type 5	0.4864	0.7100	0.4806	0.4490

Sample Range: 13-184, Sample Size: 172 (extended, incl. 34-period cycle)

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.2736	0.2794	0.3938	0.1703
Data type 2	0.2727	0.2790	0.3951	0.1789
Data type 3	0.2743	0.2802	0.3935	0.1622
Data type 4	0.2749	0.2808	0.3975	0.1964
Data type 5	0.2806	0.2748	0.4618	0.1424

The discrepancy measure is computed by eq.(60).

Table 3 Correlation with True Values: Cyclical Components

Sample Range: 13-180, Sample Size: 168	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	1.0000	1.0000	0.9997	0.9105
Data type 2	0.9993	0.9993	0.9997	0.9096
Data type 3	0.9989	0.9988	0.9996	0.9105
Data type 4	0.9969	0.9968	0.9971	0.9066
Data type 5	0.9987	0.9986	0.9243	0.9178
Sample Range: 13-184, Sample Size: 172	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9530	0.9846	0.9997	0.9188
Data type 2	0.9466	0.9772	0.9997	0.9155
Data type 3	0.9547	0.9865	0.9997	0.9214
Data type 4	0.9359	0.9647	0.9975	0.9086
Data type 5	0.9571	0.9889	0.9358	0.9307
Sample Range: 13-184, Sample Size: 172 (folded extension)	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9568	0.9031	0.9997	0.9308
Data type 2	0.9567	0.9027	0.9997	0.9259
Data type 3	0.9568	0.9034	0.9997	0.9354
Data type 4	0.9569	0.9037	0.9975	0.9182
Data type 5	0.9562	0.9015	0.9358	0.9423
Sample Range: 13-184, Sample Size: 172 (extended, incl. 34-period cycle)	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9618	0.9819	0.9996	0.9867
Data type 2	0.9621	0.9819	0.9996	0.9853
Data type 3	0.9616	0.9819	0.9995	0.9879
Data type 4	0.9615	0.9820	0.9972	0.9825
Data type 5	0.9601	0.9799	0.9346	0.9900

Table 4　Correlation with True Values: Trend Components, $\frac{\sigma_T}{\sigma_C} = 16$

Sample Range: 13-180, Sample Size: 168

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	1.0000	0.9999	0.9996	0.9993
Data type 2	1.0000	0.9999	0.9996	0.9993
Data type 3	1.0000	0.9999	0.9996	0.9993
Data type 4	1.0000	0.9999	0.9996	0.9993
Data type 5	1.0000	0.9999	0.9995	0.9994

Sample Range: 13-184, Sample Size: 172

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9989	0.9987	0.9996	0.9993
Data type 2	0.9989	0.9987	0.9996	0.9993
Data type 3	0.9989	0.9987	0.9996	0.9993
Data type 4	0.9989	0.9987	0.9996	0.9992
Data type 5	0.9990	0.9988	0.9995	0.9994

Sample Range: 13-184, Sample Size: 172 (folded extension)

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9994	0.9987	0.9996	0.9994
Data type 2	0.9994	0.9987	0.9996	0.9994
Data type 3	0.9994	0.9988	0.9996	0.9995
Data type 4	0.9994	0.9988	0.9996	0.9994
Data type 5	0.9995	0.9989	0.9995	0.9996

Sample Range: 13-184, Sample Size: 172 (extended, incl. 34-period cycle)

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9998	0.9998	0.9996	0.9999
Data type 2	0.9998	0.9998	0.9996	0.9999
Data type 3	0.9998	0.9998	0.9996	0.9999
Data type 4	0.9998	0.9998	0.9996	0.9999
Data type 5	0.9998	0.9998	0.9995	1.0000

Table 5 Correlation with True Values: Trend Components, $\frac{\sigma_T}{\sigma_C} = 2$

Sample Range: 13-180, Sample Size: 168

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9992	0.9920	0.9745	0.9613
Data type 2	0.9992	0.9920	0.9745	0.9612
Data type 3	0.9992	0.9921	0.9746	0.9614
Data type 4	0.9992	0.9921	0.9749	0.9615
Data type 5	0.9994	0.9931	0.9767	0.9649

Sample Range: 13-184, Sample Size: 172

	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9377	0.9261	0.9730	0.9586
Data type 2	0.9380	0.9265	0.9730	0.9587
Data type 3	0.9377	0.9261	0.9731	0.9586
Data type 4	0.9400	0.9289	0.9738	0.9596
Data type 5	0.9443	0.9337	0.9757	0.9629

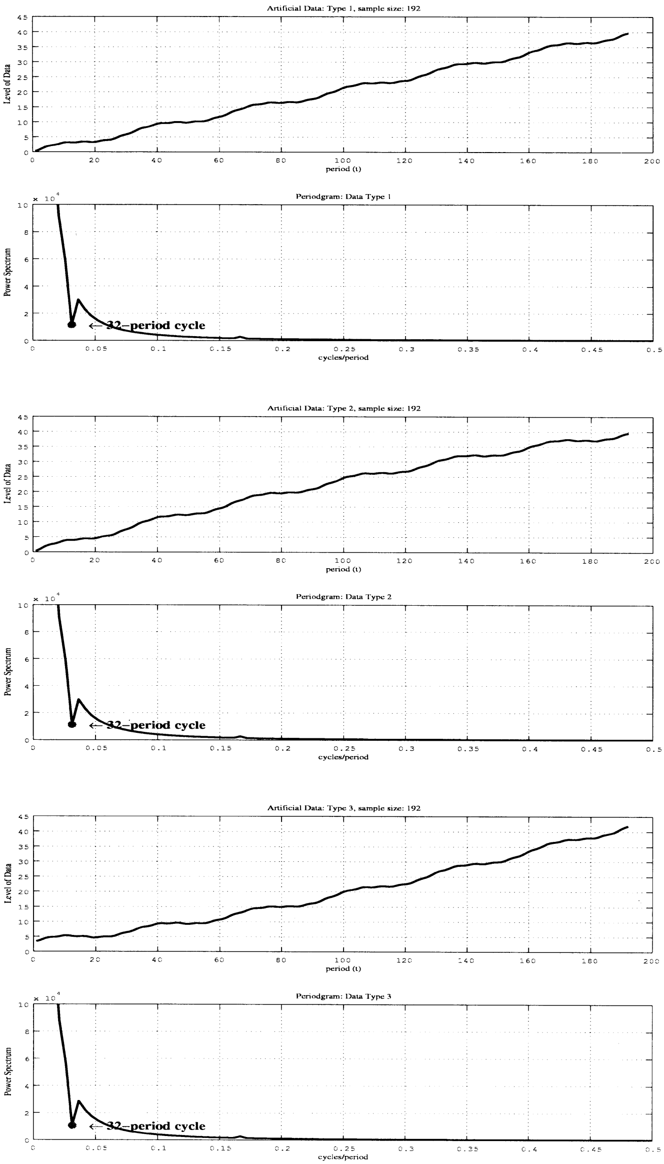
Sample Range: 13-184, Sample Size: 172 (folded extension)

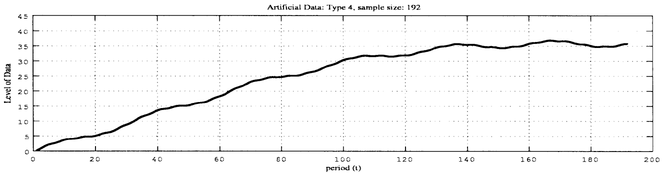
	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9638	0.9303	0.9730	0.9666
Data type 2	0.9638	0.9304	0.9730	0.9664
Data type 3	0.9639	0.9305	0.9731	0.9670
Data type 4	0.9648	0.9322	0.9738	0.9666
Data type 5	0.9680	0.9375	0.9757	0.9705

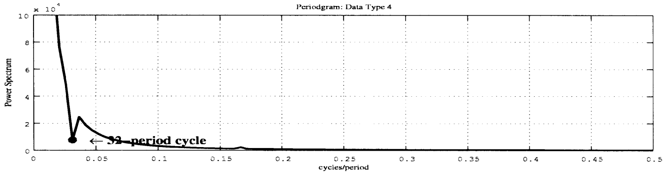
Sample Range: 13-184, Sample Size: 172 (extended, incl. 34-period cycle)

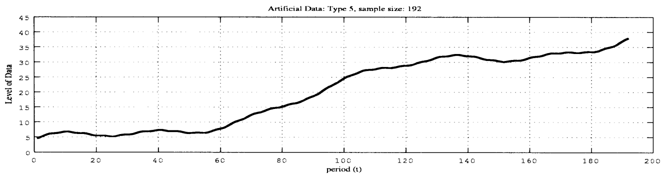
	Butterworth (frequency)	Hamming-Windowed	Baxter-King Filter	CF (RW) Filter
Data type 1	0.9879	0.9876	0.9764	0.9954
Data type 2	0.9879	0.9876	0.9764	0.9953
Data type 3	0.9880	0.9877	0.9764	0.9955
Data type 4	0.9882	0.9880	0.9770	0.9953
Data type 5	0.9894	0.9892	0.9787	0.9961

Fig. 1　Artificial Data: Type 1 - Type 5




Artificial Data: Type 4, sample size: 192
Level of Data
period (t)


Periodgram: Data Type 4
x 10^4
Power Spectrum
← 32-period cycle
cycles/period


Artificial Data: Type 5, sample size: 192
Level of Data
period (t)

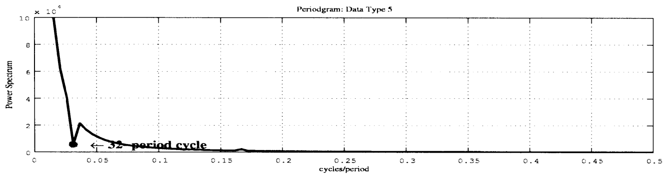
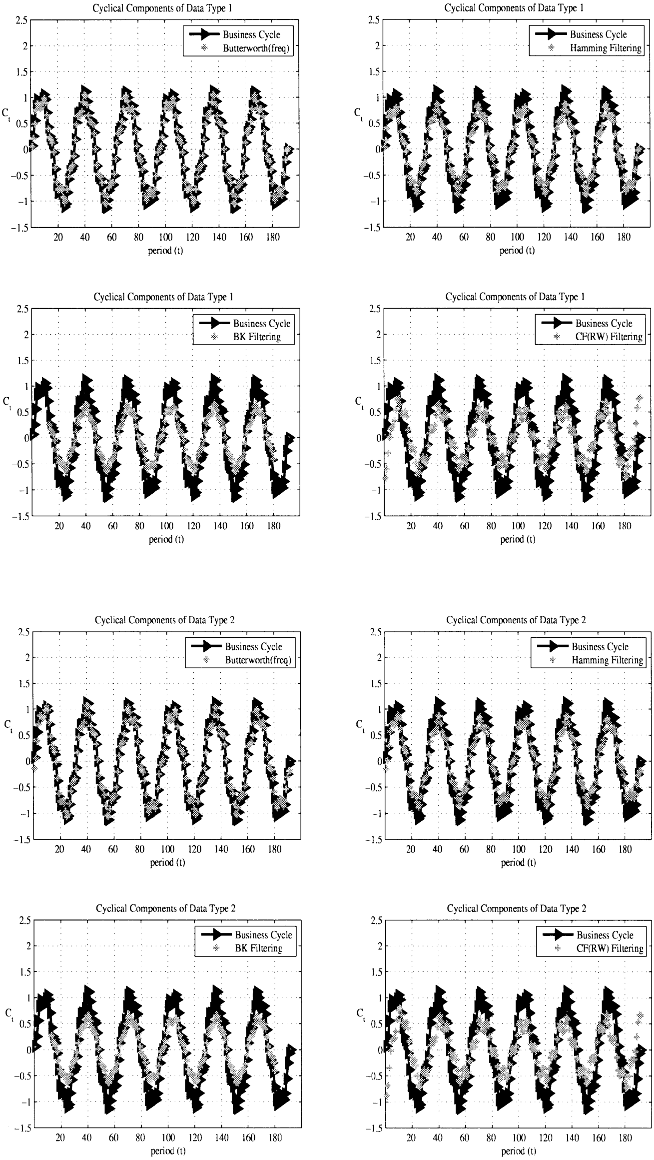
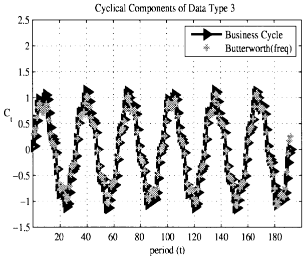
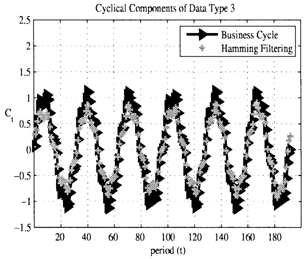
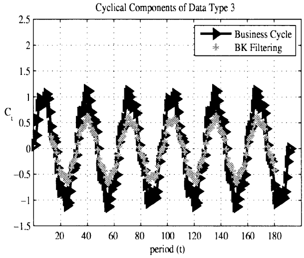

Periodgram: Data Type 5
x 10^4
Power Spectrum
← 32-period cycle
cycles/period

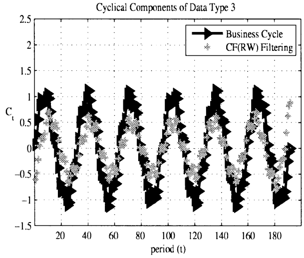
Fig. 2　Phase Shifts: 192 samples

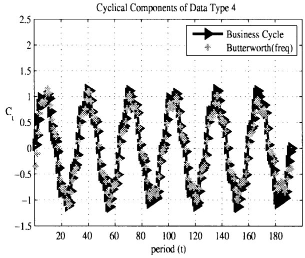


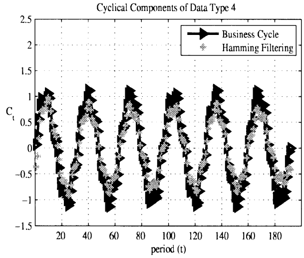

Cyclical Components of Data Type 3
Business Cycle
Butterworth(freq)
C_t
period (t)

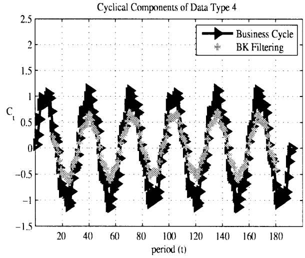

Cyclical Components of Data Type 3
Business Cycle
Hamming Filtering
C_t
period (t)

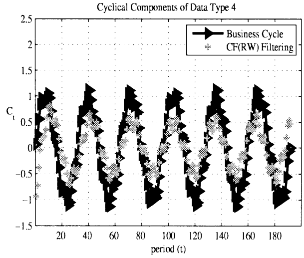

Cyclical Components of Data Type 3
Business Cycle
BK Filtering
C_t
period (t)

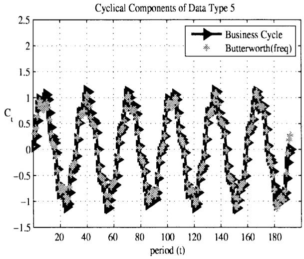

Cyclical Components of Data Type 3
Business Cycle
CF(RW) Filtering
C_t
period (t)

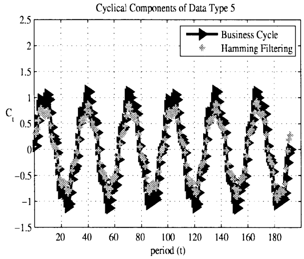

Cyclical Components of Data Type 4
Business Cycle
Butterworth(freq)
C_t
period (t)

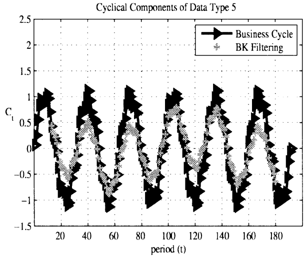

Cyclical Components of Data Type 4
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 4
Business Cycle
BK Filtering
C_t
period (t)


Cyclical Components of Data Type 4
Business Cycle
CF(RW) Filtering
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
Butterworth(freq)
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
BK Filtering
C_t
period (t)

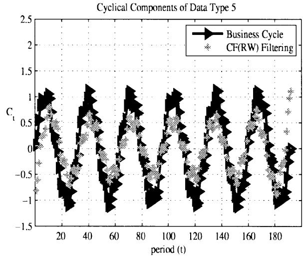
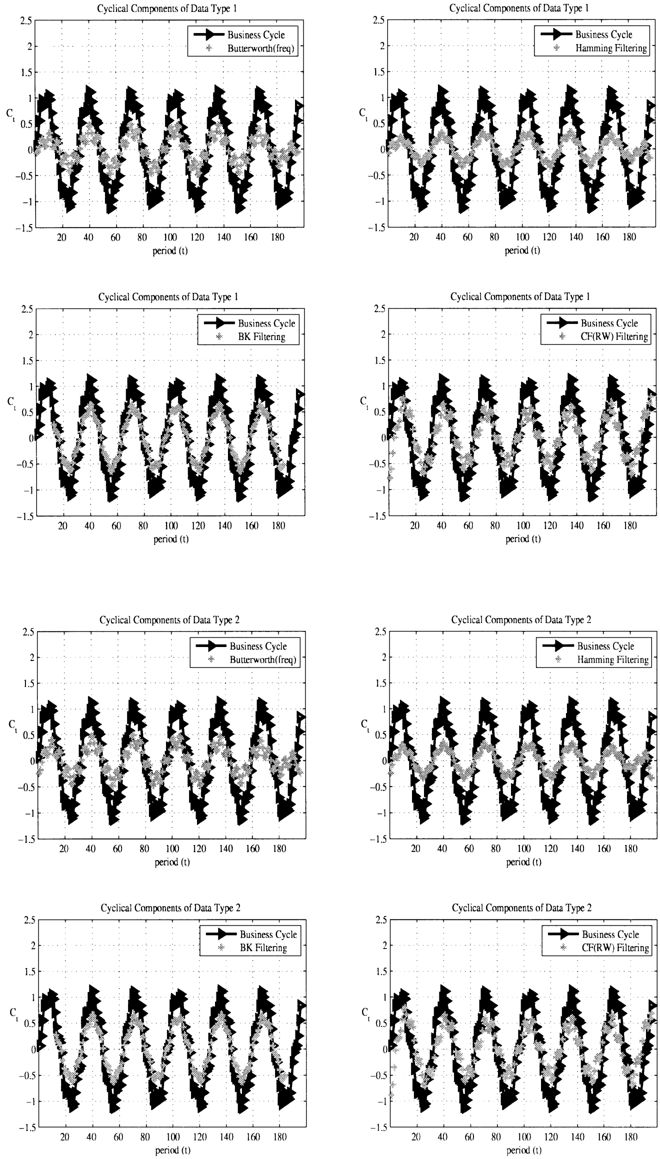
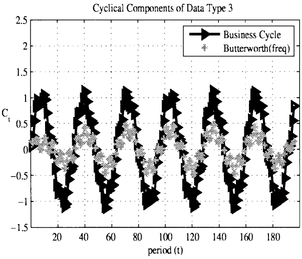
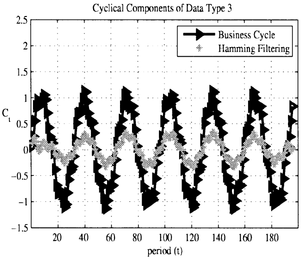
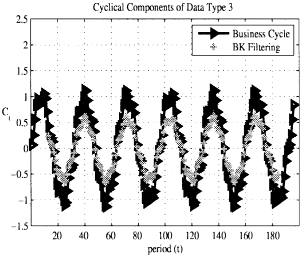

Cyclical Components of Data Type 5
Business Cycle
CF(RW) Filtering
C_t
period (t)

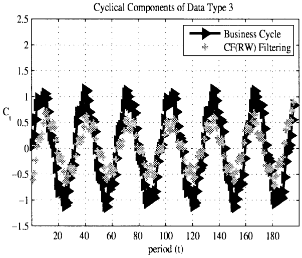
Fig. 3 Phase Shifts: 196 samples

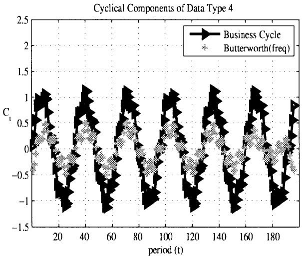


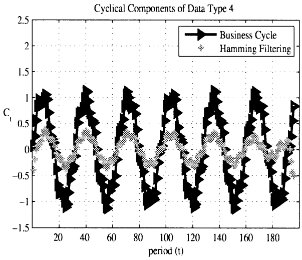

Cyclical Components of Data Type 3
Business Cycle
Butterworth(freq)
C_t
period (t)

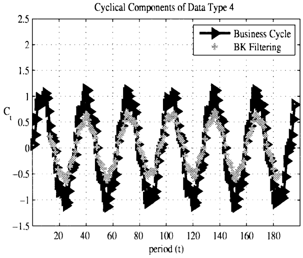

Cyclical Components of Data Type 3
Business Cycle
Hamming Filtering
C_t
period (t)

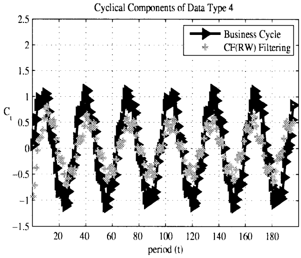

Cyclical Components of Data Type 3
Business Cycle
BK Filtering
C_t
period (t)

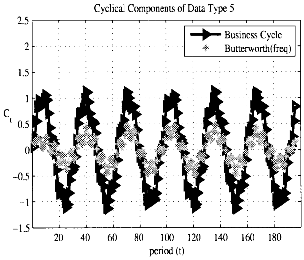

Cyclical Components of Data Type 3
Business Cycle
CF(RW) Filtering
C_t
period (t)

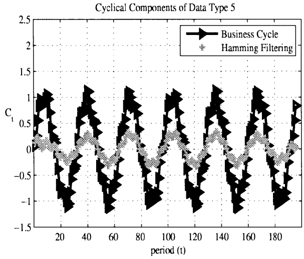

Cyclical Components of Data Type 4
Business Cycle
Butterworth(freq)
C_t
period (t)

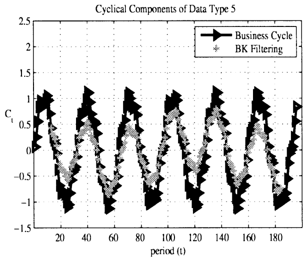

Cyclical Components of Data Type 4
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 4
Business Cycle
BK Filtering
C_t
period (t)


Cyclical Components of Data Type 4
Business Cycle
CF(RW) Filtering
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
Butterworth(freq)
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 5
Business Cycle
BK Filtering
C_t
period (t)

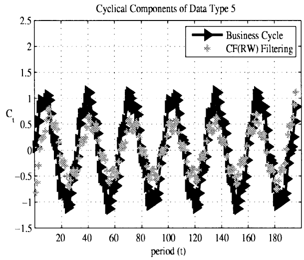
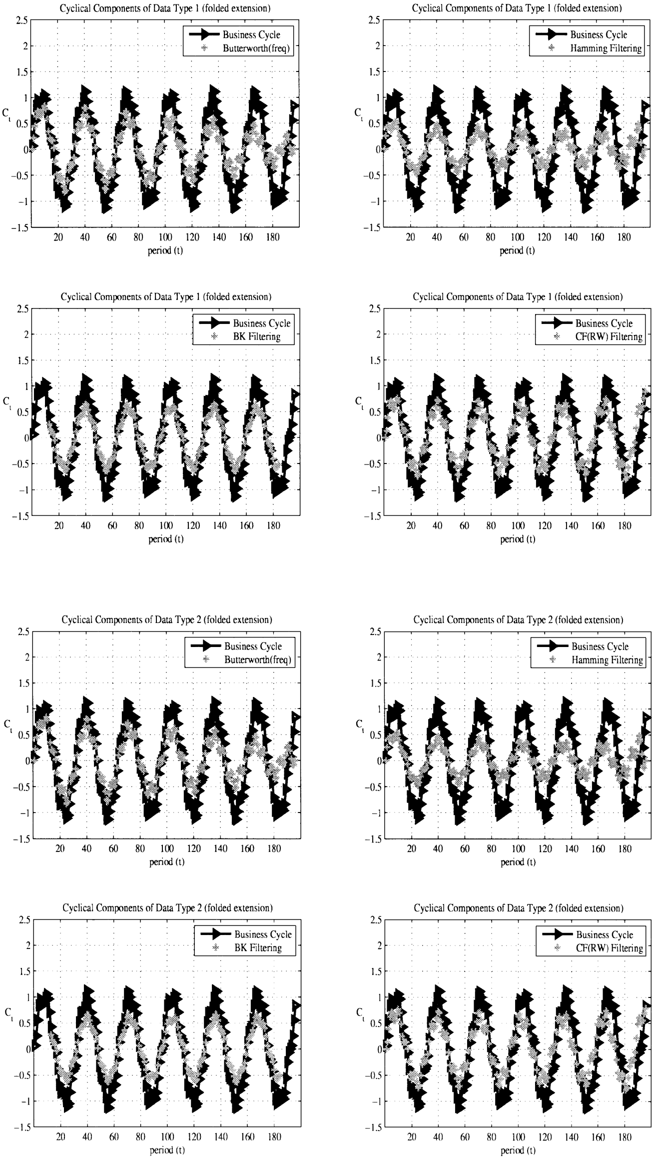
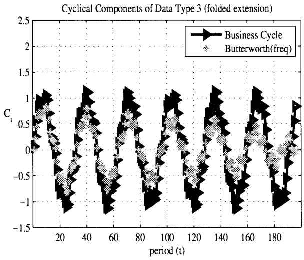
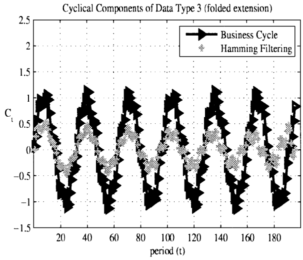
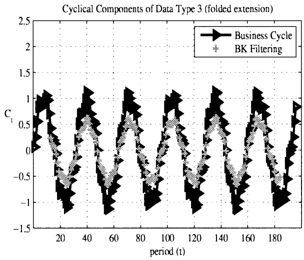

Cyclical Components of Data Type 5
Business Cycle
CF(RW) Filtering
C_t
period (t)

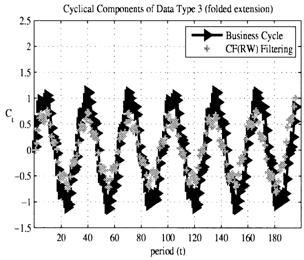
Fig. 4　Phase Shifts: 196 samples (folded extension)

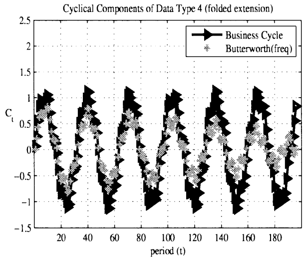


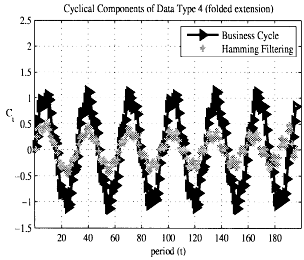

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)

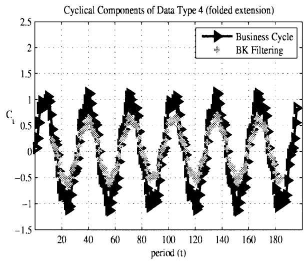

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)

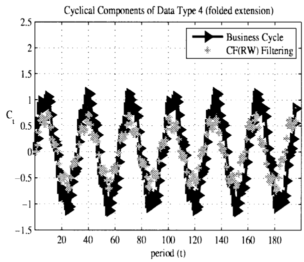

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)

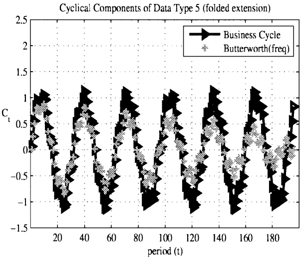

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)

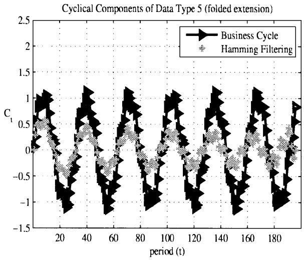

Cyclical Components of Data Type 4 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)

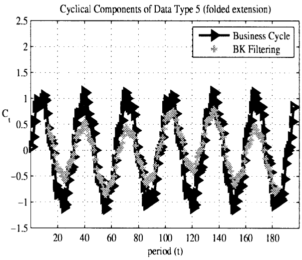

Cyclical Components of Data Type 4 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 4 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)


Cyclical Components of Data Type 4 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)

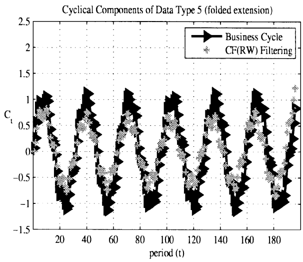
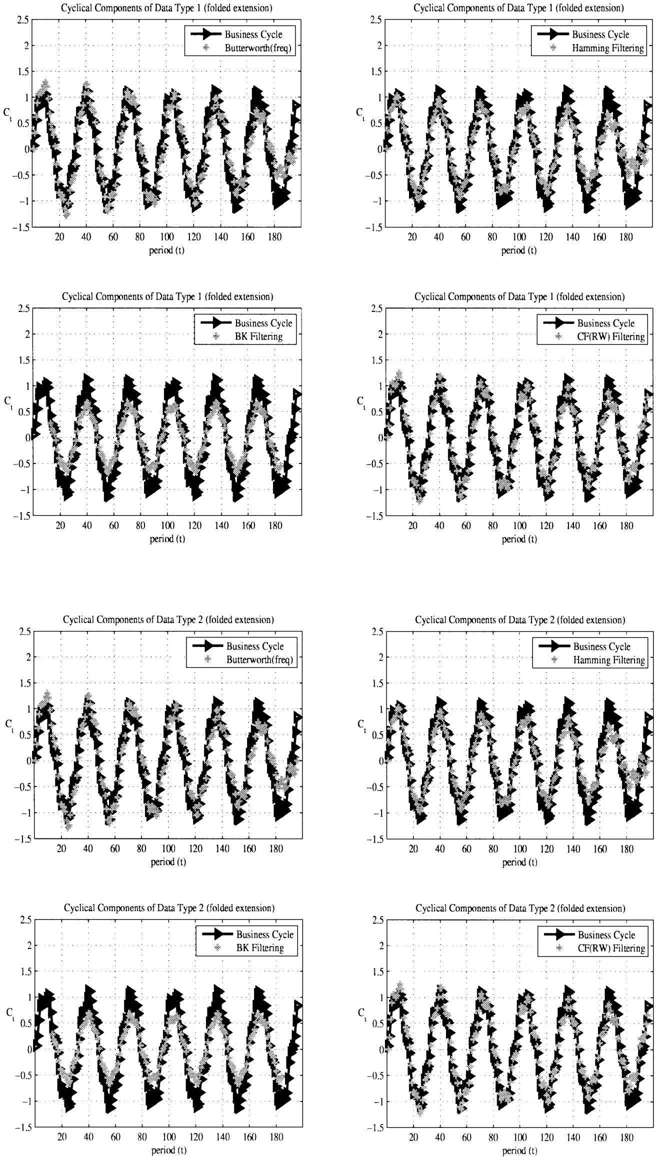
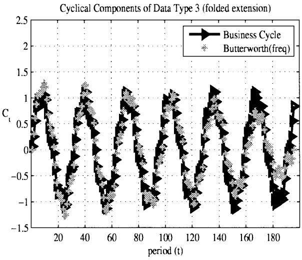
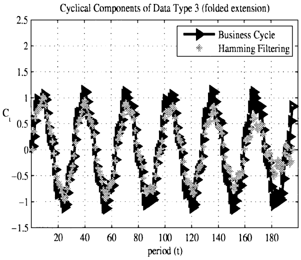
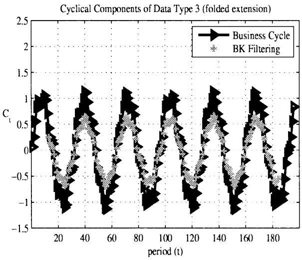

Cyclical Components of Data Type 5 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)

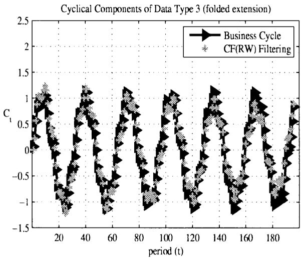
Fig. 5 Phase Shifts: 196 samples (incl. 34-period cycle)

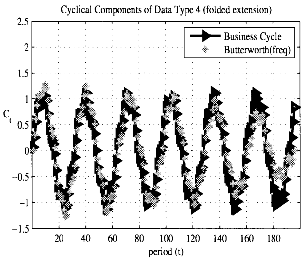


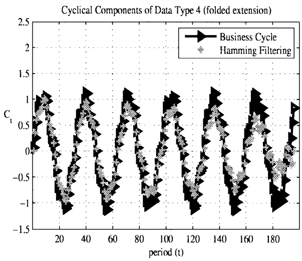

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)

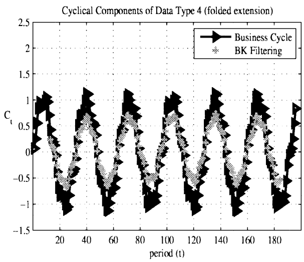

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)

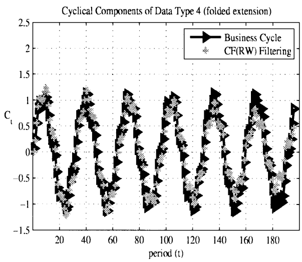

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)

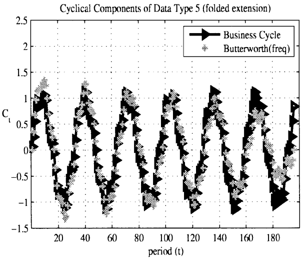

Cyclical Components of Data Type 3 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)

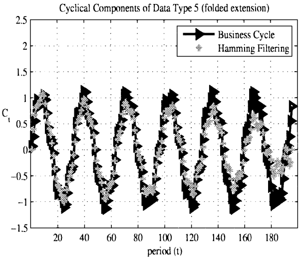

Cyclical Components of Data Type 4 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)

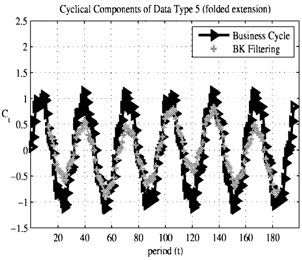

Cyclical Components of Data Type 4 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 4 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)


Cyclical Components of Data Type 4 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
Butterworth(freq)
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
Hamming Filtering
C_t
period (t)


Cyclical Components of Data Type 5 (folded extension)
Business Cycle
BK Filtering
C_t
period (t)

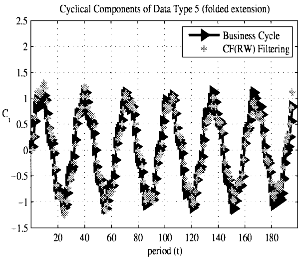

Cyclical Components of Data Type 5 (folded extension)
Business Cycle
CF(RW) Filtering
C_t
period (t)

Fig. 6 Trend Estimates: Data Type 1, 192 samples, $\frac{\sigma_T}{\sigma_C} = 16$

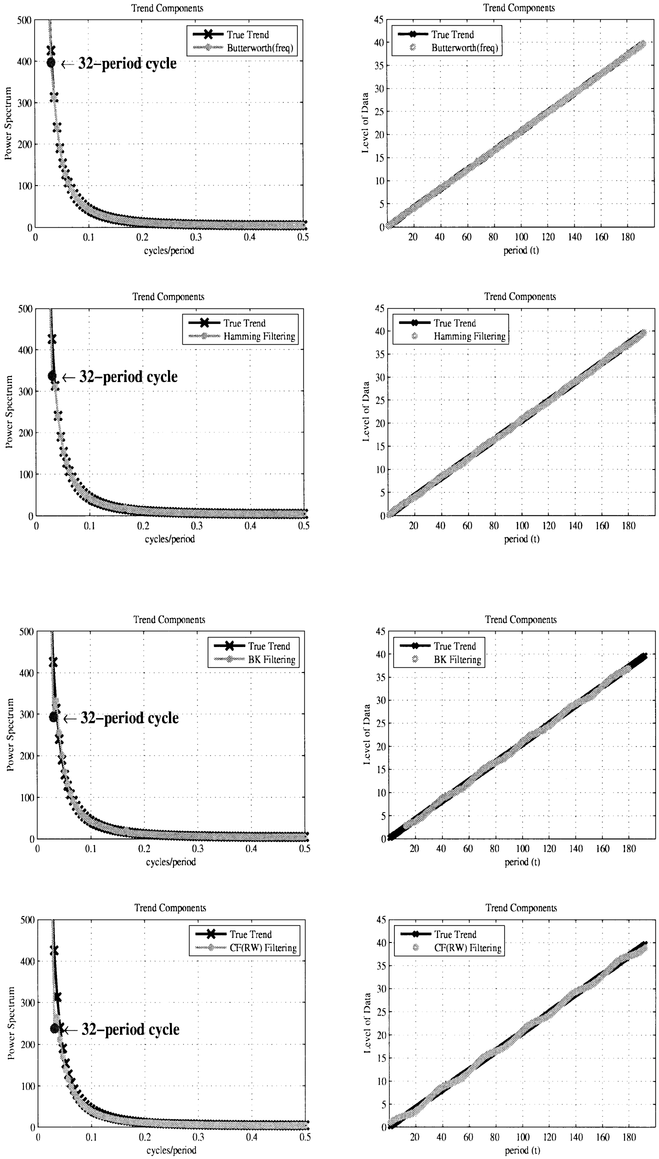


Fig. 7 Trend Estimates: Data Type 4, 192 samples, $\frac{\sigma_T}{\sigma_C} = 16$

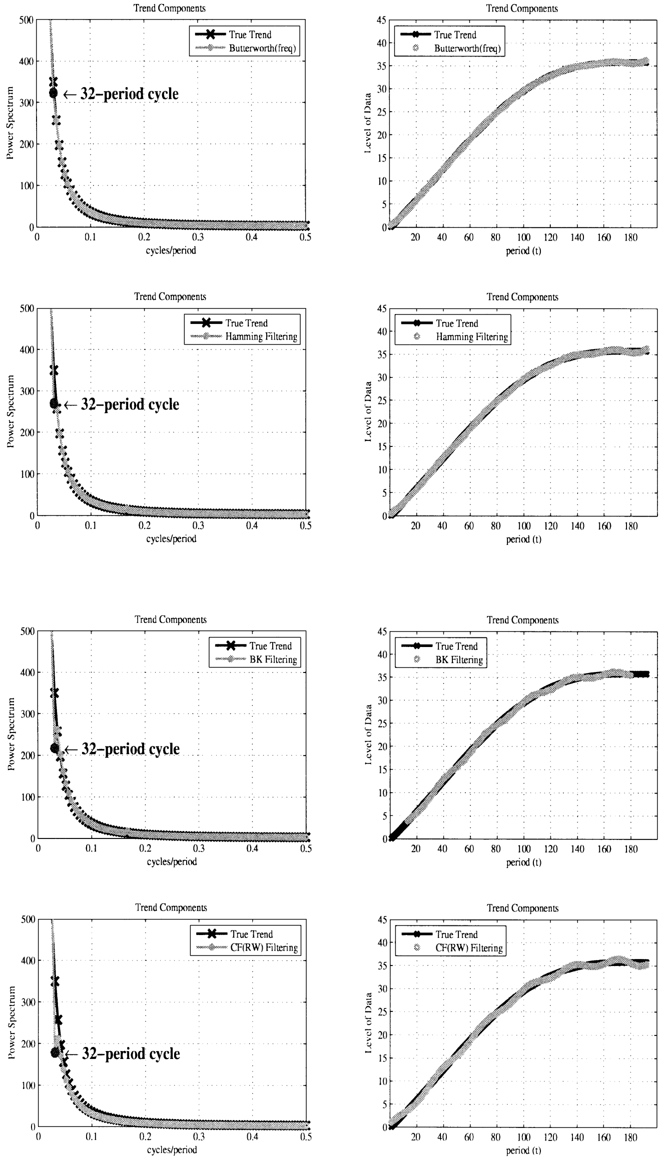


Fig. 8　Trend Estimates: Data Type 1, 196 samples, $\frac{\sigma_T}{\sigma_C} = 16$

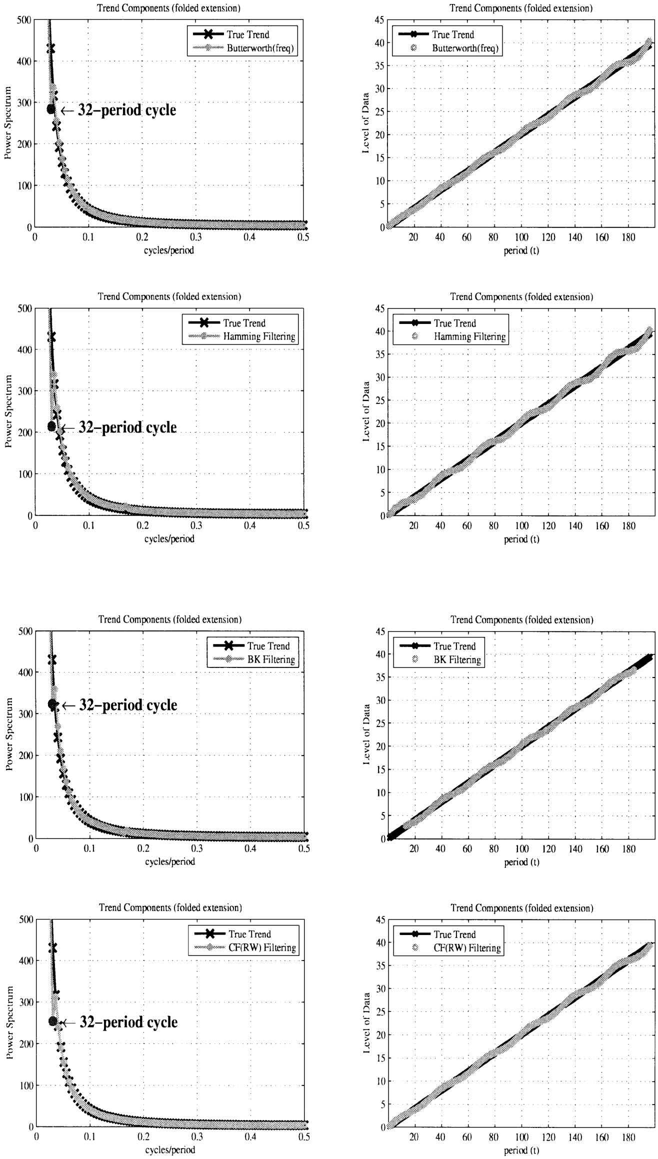


Fig. 9 Trend Estimates: Data Type 4, 196 samples, $\frac{\sigma_T}{\sigma_C} = 16$

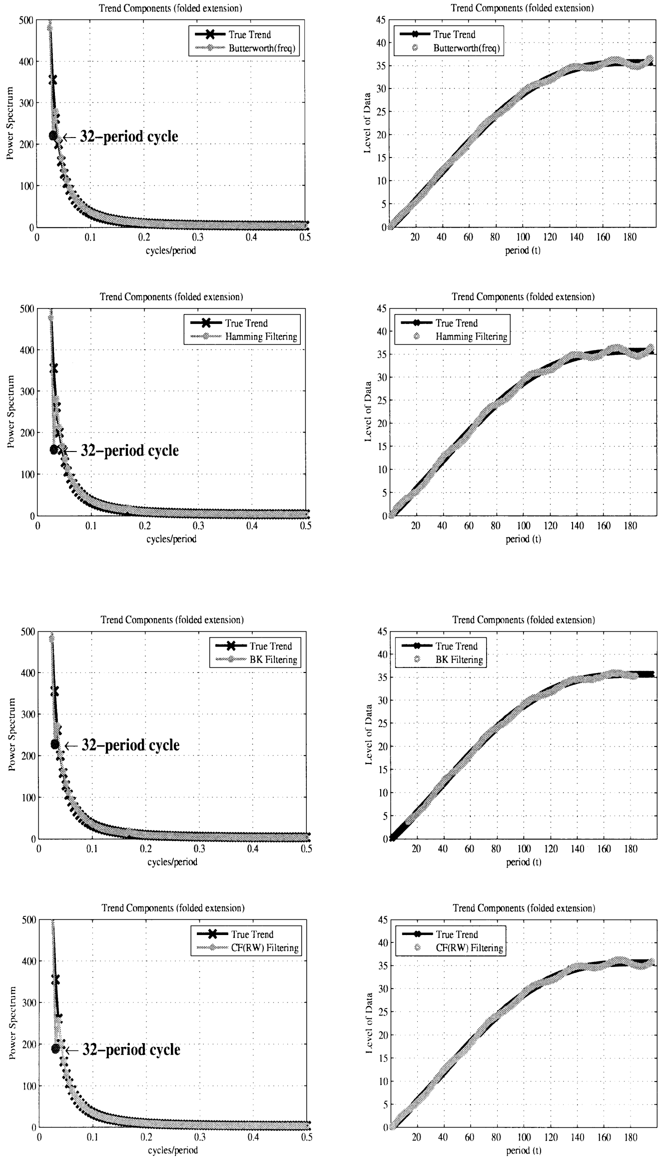


Fig. 10　Artificial Data: Type 1 and Type 4, 192 samples, $\frac{\sigma_T}{\sigma_C} = 2$

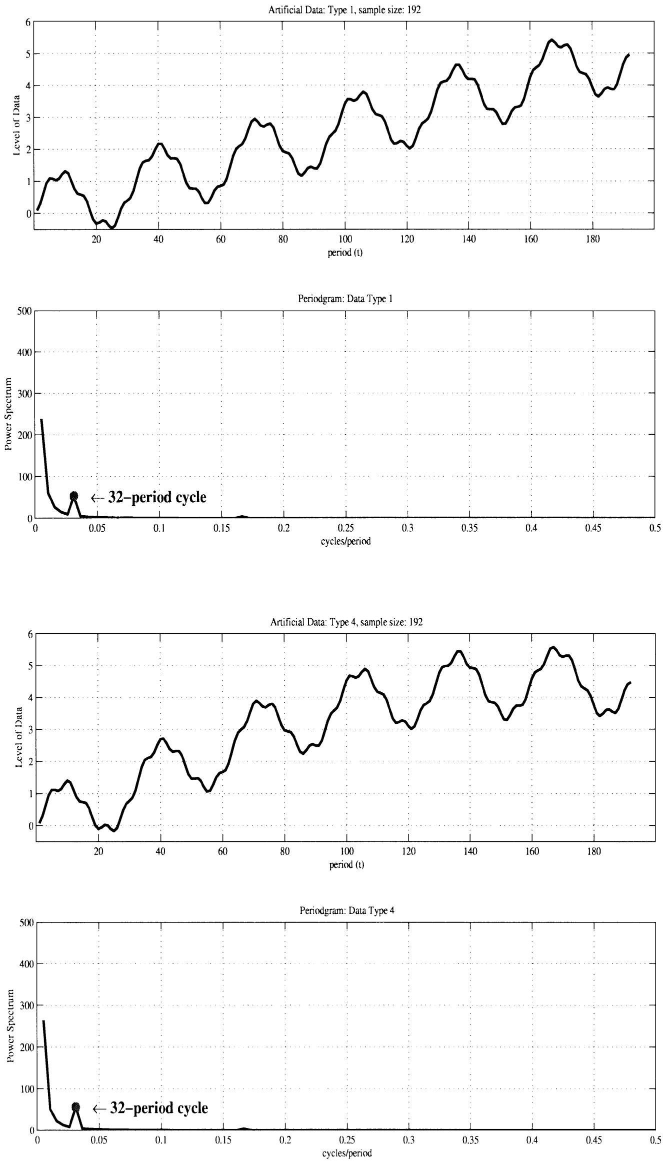


Fig. 11 Trend Estimates: Data Type 1, 192 samples, $\frac{\sigma_T}{\sigma_C} = 2$

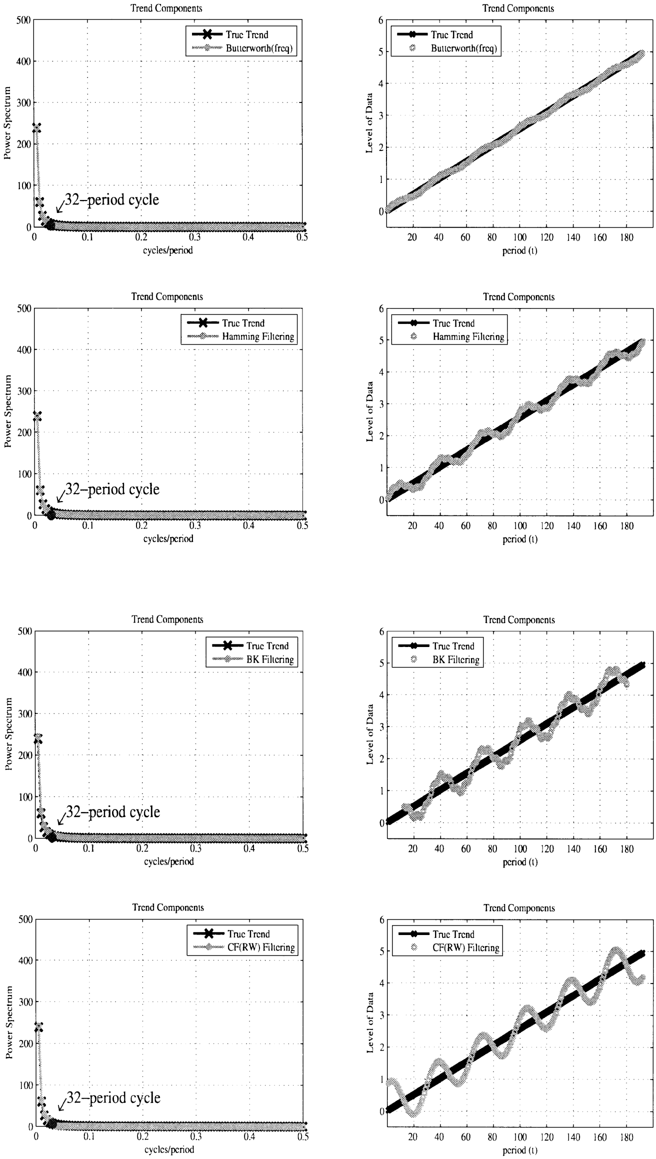


Fig. 12　Trend Estimates: Data Type 4, 192 samples, $\frac{\sigma_T}{\sigma_C} = 2$

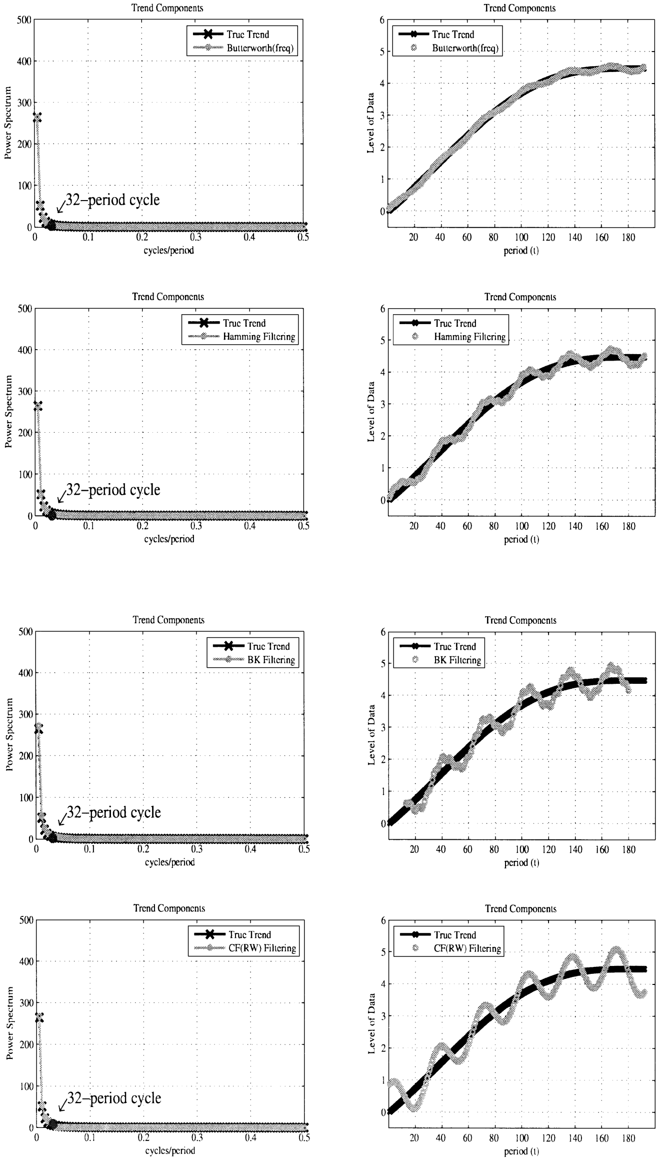


Fig. 13 Trend Estimates: Data Type 1, 196 samples, $\frac{\sigma_T}{\sigma_C} = 2$

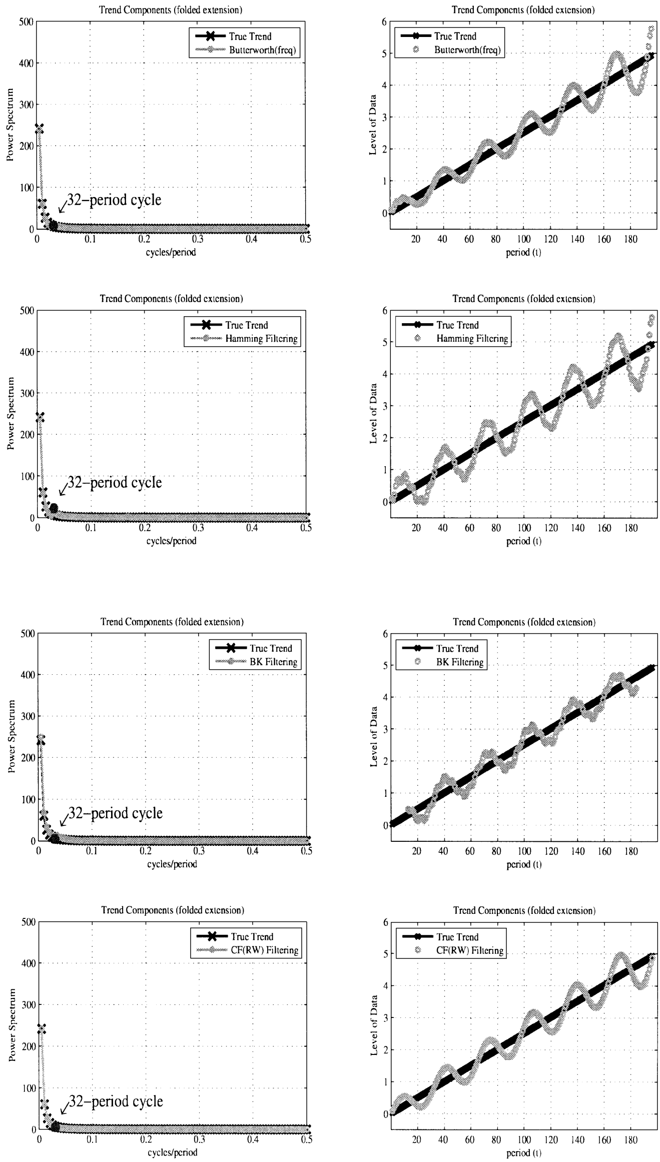


Fig. 14　Trend Estimates: Data Type 4, 196 samples, $\frac{\sigma_T}{\sigma_C} = 2$

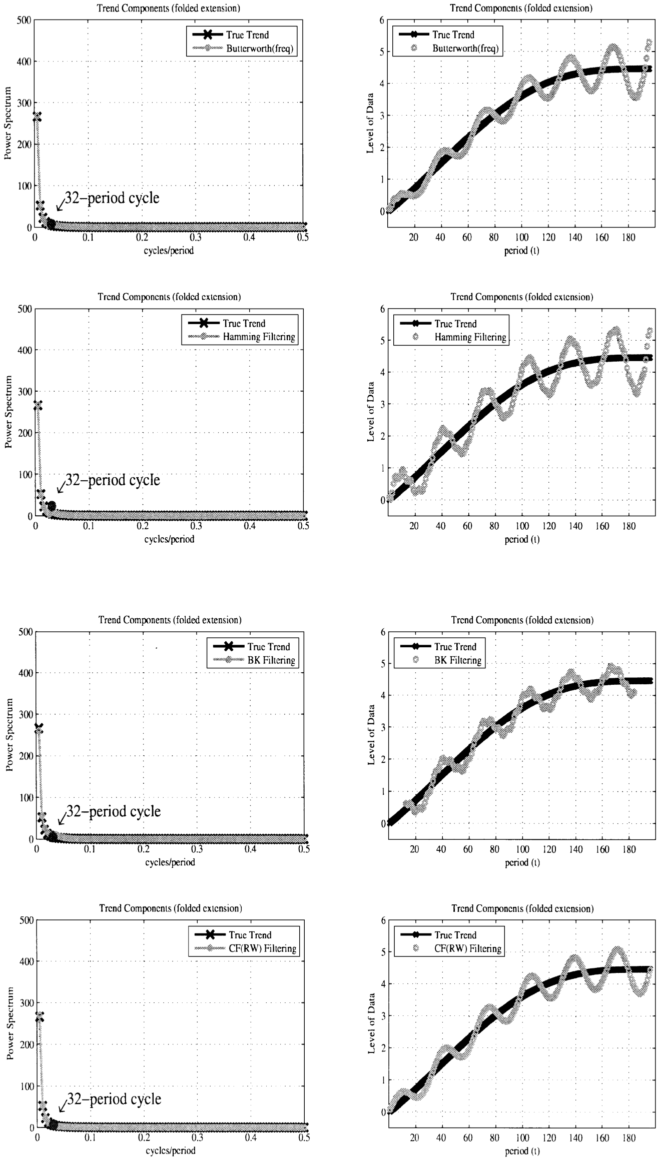


Fig. 15 Trend Estimates: Data Type 1, 196 samples, $\frac{\sigma_T}{\sigma_C} = 2$ (incl. 34-period cycle)

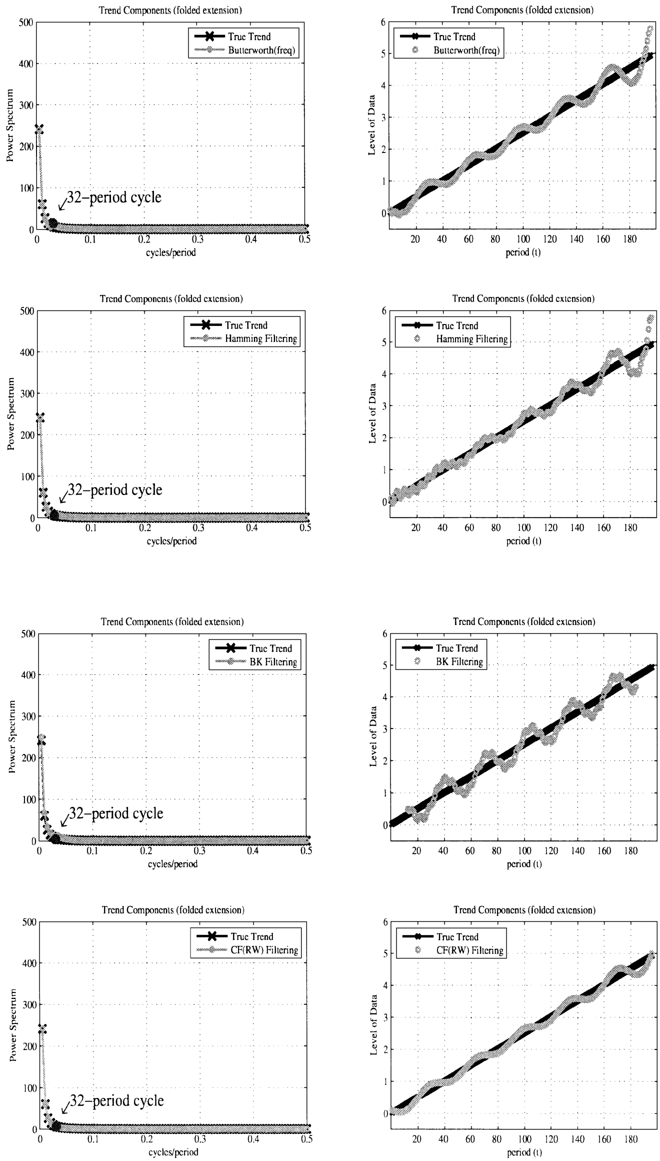


Fig. 16　Trend Estimates: Data Type 4, 196 samples, $\frac{\sigma_T}{\sigma_C} = 2$ (incl. 34-period cycle)