A Multivariate Band-Pass Filter

Valle e Azevedo, João

02. January 2008

A Multivariate Band-Pass Filter

João Valle e Azevedo*

Banco de Portugal & Universidade NOVA de Lisboa

January 2, 2008

Abstract

We develop a multivariate filter which is an optimal (in the mean squared error sense) approximation to the ideal filter that isolates a specified range of fluctuations in a time series, e.g., business cycle fluctuations in macroeconomic time series. This requires knowledge of the true second-order moments of the data. Otherwise these can be estimated and we show empirically that the method still leads to relevant improvements of the extracted signal, especially in the endpoints of the sample. Our filter is an extension of the univariate filter developed by Christiano and Fitzgerald (2003). Specifically, we allow an arbitrary number of covariates to be employed in the estimation of the signal.

We illustrate the application of the filter by constructing a business cycle indicator for the U.S. economy. The filter can additionally be used in any similar signal extraction problem demanding accurate real-time estimates.

JEL Classification: C14, C32, E32

Keywords: Band-Pass filter, Business Cycle indicator, Revisions

*Address: Av. Almirante Reis, 71-6th floor, 1150-012 Lisboa, Portugal; E-mail: jvazevedo@bportugal.pt; Phone: +351 213130163. Most research was done while I was a graduate student at Stanford University. I gratefully acknowledge the Portuguese Government, through the Fundação para a Ciência e Tecnologia, for financial support during my graduate studies.

1 Introduction

Undertaking fiscal and monetary policies requires knowledge about the state of the economy. Given the mixed signals provided by the various macroeconomic time series available, the task of accurately determining this state is a challenging one. Specifically, it is hard to determine precisely in real-time which movements in economic activity are part of a slowly evolving stochastic trend and which movements are attributable to the typical business cycle fluctuations. This is true even if everyone agrees on a statistical definition of business cycle fluctuations. A popular definition uses concepts from spectral analysis to define business cycle fluctuations in macroeconomic time series as "fluctuations with a specified range of periodicities" (Baxter and King 1999). The periodicities typically range from 6 to 32 quarters. Various methods have been employed to isolate the desired periodicities (or frequencies) in the data. These amount to applying band-pass filters to the series of interest and seem to be sufficient if the purpose of the analyst is to look at historical or simulated data. However, their real-time performance leaves much to be desired.

Our main contribution is the development of an approximation to the business cycle fluctuations, defined as above, that incorporates information from an arbitrary number of time series. All the existing methods are univariate. We explore how additional information can reduce the uncertainty associated with real-time estimates of business cycle fluctuations. We will show empirically and through a simulation exercise that under certain conditions on the relations between the covariates and the series of interest, our method leads to improvements of the extracted signal, meaning that subsequent revisions of real-time estimates are the smallest among the methods analyzed. Still, if these conditions are not met we find no deterioration of that signal. This means essentially that if the covariates are not helpful in forecasting the dynamics of the series of interest they will be given a negligible weight in the determination of the filtered series. The main application of this method is in the construction of business cycle indicators, although it can be used in any other similar signal extraction problem demanding precise real-time estimates. Specifically, our results can straightforwardly be adapted to produce optimal approximations to any (absolutely summable and stationary) distributed lag of the series of interest. This includes, for instance, real-time approximations to non-parametric seasonal adjustment filters.

Baxter and King (1999) were the first to provide a criterial method to isolate business cycle fluctuations. Their approach results in a symmetric filter (BK filter) that does not depend on the data generating process (DGP) of the series being filtered. However, the BK filters cannot be used in real-time since observations are lost in the endpoints of the sample. This is not the case with the Hodrick-Prescott (HP) filter (see Hodrick and Prescott 1997), which is therefore more related to our approach as well as to our objectives. The HP filter has long been used to eliminate low frequencies in the data. The fact that the HP filter is indeed a high-pass filter (a filter that eliminates only low frequencies and retains without distortion high frequencies) was pointed by King and Rebelo (1993). Using the HP filter, it is easy to construct a band-pass filter, by applying successively an HP high-pass filter and the complementary of another HP high-pass filter (a low-pass filter). The choice of the smoothing parameters can also be reconciled with the definition of business cycle fluctuations as fluctuations with a specified range of periodicities (see Pedersen 2001). Butterworth filters, which can be seen as generalizations of the HP filter, provide better approximations (at least in infinite samples) to the ideal filter, the filter that would perfectly isolate the desired frequencies (see Gomez 2001). Although real-time estimates can be obtained, the major problem with these filters is their behavior in the endpoints of the sample. Even though improvements can arise if the series are extended with backcasts and forecasts, the method is not unifying, that is, there is no attempt to approximate the ideal filter and to use the information from the DGP simultaneously.

A major development comes with the work of Christiano and Fitzgerald (2003). They provided a solution to the endpoints problem, by developing a band-pass filter (henceforth CF filter) which is optimal (in the mean squared error sense) for every observation in the sample (and obviously and most importantly for the endpoints), given that the true DGP is known. The filter is optimal in the class of filters that uses only the series of interest to determine an estimate of the exact band-pass filtered series. Christiano and Fitzgerald (2003) show that even if the DGP is not exactly known the approximation is quite reasonable. However, it is clear that there will be always revisions of the estimates once new data is available, especially in the endpoints of the sample. By solving a problem very similar to the one in Christiano and Fitzgerald (2003), but where an arbitrary number of covariates can additionally be used to determine the estimates of the ideal filtered series at any point in time, we are able to further reduce the revisions of the estimates in the endpoints. While optimal over the entire sample, our multivariate band-pass filter achieves small gains in the middle of the sample, where the univariate method (and even the BK filter) are extremely accurate. But again, it can significantly help in the improvement of the signal in the endpoints of the sample. If the covariates used are highly correlated with the series of interest and if they are good predictors of the dynamics of this series, it is likely that they will be useful in determining accurately the cyclical position at any point.

Our approach, as well as those referred above, can be regarded as non-parametric. Model-based (or parametric) methods have also been used to construct business cycle indicators. Harvey and Trimbur (2003) propose structural models for which the extraction of a cycle component is equivalent to using a band-pass filter. Using the components in Harvey and Trimbur (2003) and incorporating an extension by Rünstler (2004) that allows for phase shifts in the cyclical components of multiple time series Valle e Azevedo, Koopman and Rua (2006) construct a business cycle indicator which can be seen as a multivariate band-pass filter. Although the filter leads to good real-time properties, it does not aim at approximating an ideal filter isolating a pre-defined range of frequencies. Such is the aim of this paper.

The remainder of the paper is organized as follows. In section 2 we set out the problem to be solved and present the solution. In section 3 we look at the properties of the derived filter when good indicators are available. In section 4 we compute a business cycle indicator, obtained by applying the filter to U.S. GDP using a moderate set of available indicators, and compare its performance with other filtering methods. Section 5 concludes.

2 Multivariate Band-Pass filtering

2.1 Spectral representation

Let $\{X_t\}$ be a covariance-stationary vector sequence with mean zero and define the spectrum of $\{X_t\}$ as:

$$S_X(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{-i\omega k}\Gamma(k), \quad -\pi \leq \omega \leq \pi$$

where $i^2 = -1$, ω denotes the frequency measured in radians and $\Gamma(k)$ is the autocovariance matrix of $\{X_t\}$ at lag k. It is well known (see, e.g., Brockwell and Davis 1991, p. 456) that there exists a right-continuous orthogonal increment process $\{Z(\omega), -\pi \leq \omega \leq \pi\}$ such that:

$$i)\ E[(Z(\omega) - Z(-\pi))(Z(\omega) - Z(-\pi))^*] = F_X(\omega) = \int_{-\pi}^{\omega} S_X(\omega)d\omega, \quad -\pi \leq \omega \leq \pi$$

where $*$ denotes conjugate transposed and

$$ii)\ X_t = \int_{-\pi}^{\pi} e^{i\omega t}dZ(\omega) \ \ a.e.$$

So, X_t can be decomposed into an infinite weighted sum of orthogonal fluctuations, each with frequency ω. $S_X(\omega)$ can be interpreted as the decomposition of the variance of X_t in terms of these fluctuations. $S_X(\omega)$ contains the same information as the second order moments characterized by $\Gamma(k)$, $k = 0, \pm1, \pm2, \ldots.$ $S_X(\omega)$ and $\Gamma(k)$ form a pair of Fourier transforms in that:

$$\Gamma(k) = \int_{-\pi}^{\pi} e^{i\omega k}dF_X(\omega) = \int_{-\pi}^{\pi} e^{i\omega k}S_X(\omega)d\omega, k = 0, \pm1, \pm2, \ldots \tag{1}$$

If we apply a time-invariant linear filter $H(L) = \sum_{j=-\infty}^{\infty} H_j L^j$ where $L^j X_t = X_{t-j}$ and such that $\sum_{j=-\infty}^{\infty} |H_j| < \infty$ to the sequence $\{X_t\}$ we obtain a filtered sequence $Y_t = \sum_{k=-\infty}^{\infty} H_j X_{t-j}$. It is easy to verify that the spectrum of $\{Y_t\}$ is given by:

$$S_Y(\omega) = H(e^{-i\omega}) S_X(\omega) H^{'}(e^{i\omega}) \tag{2}$$

Now suppose some elements of $\{X_t\}$ have one or more unit roots. A relation just like (2) holds if the filter $H(L)$ renders $\{X_t\}$ stationary, given that we define $S_X(\omega)$ as the multivariate pseudo-spectrum of $\{X_t\}$. The pseudo-spectrum can be viewed as the limit of the spectrum of a covariance stationary process when the smallest autoregressive roots converge to 1. Although this function has been previously defined by, e.g., Harvey (1993), Hurvich and Ray (1995) and Velasco (1999), it has only recently been given a rigorous frequency domain interpretation (as a distribution of the infinite variance over frequencies) by Bujosa, Bujosa and García-Ferrer (2002). They extend the classical spectral analysis by developing an extended Fourier transform to the field of fractions of polynomials. A pseudo-autocovariance generating function is defined to account for the presence of unit roots and the corresponding extended Fourier transform is defined as the pseudo-spectrum. The pseudo-spectrum collapses to the standard spectrum when no non-stationary roots are present, since the extended Fourier transform is just the classical Fourier transform in that case. This definition implies that the we can interpret the effects of filtering, summarized by the transfer function $H(e^{-i\omega})$, exactly as in the stationary case.

2.2 The problem and its solution

Isolating perfectly fluctuations within a range of frequencies in the spectrum of a univariate time series $\{x_t\}$ can be achieved by applying an "ideal" filter to $\{x_t\}$. Suppose we are interested in isolating the interval of frequencies $]\omega_l, \omega_h[\subset [0, \pi]$[1], corresponding to the interval of periodicities $]2\pi/\omega_h, 2\pi/\omega_l[$. The ideal filter is a linear filter with transfer function that we denote by $B(e^{-i\omega})$ in the range $[0, \pi]$. We have $B(e^{-i\omega}) = 1$ for $\omega \in]\omega_l, \omega_h[$ and 0 otherwise. That is, the ideal filter completely eliminates fluctuations with frequencies outside the band of interest and retains

[1] We restrict hereafter the analysis to the interval $[0, \pi]$, due to the symmetry of the spectrum around $\omega = 0$ in the case of real time series and real filter weights.

without distortion the remaining fluctuations. The ideally filtered series at time t is given by $y_t = B(L)x_t$, with $B(L) = \sum_{j=-\infty}^{\infty} B_j L^j$. The filter weights are well known and given by:

$$B_o = \frac{\omega_h - \omega_l}{\pi}, \qquad B_j = \frac{\sin[\omega_h j] - \sin[\omega_l j]}{\pi j}, |j| \geq 1 \tag{3}$$

Since we have an infinite number of weights in the ideal filter, we need an infinite amount of data to compute the ideally filtered series. Some sort of approximation is needed in practice. Baxter and King (1999) were the first to provide a criterion to get this approximation. The criterion is to minimize the distance between the transfer function of an applicable filter and that of the ideal band-pass filter. However, they restrict the analysis to symmetric filters and the criterion gives equal weight to the referred distance at every frequency. The use of symmetric filters disregards information from some observations in the series and the criterion does not take into account that the variance attributable to the various frequencies in a series is not in general a constant function of the frequencies. If the power of the series is concentrated in some range of frequencies a more reasonable criterion would give more weight to the distance at those frequencies. Christiano and Fitzgerald (2003) do exactly this. They give the minimum mean squared error solution to the problem of approximating the ideally filtered series by an applicable filter that can be a function of all the data points in a series. It is shown that the frequency domain version of this problem amounts to minimize a distance between the transfer function of the ideal filter and that of an applicable filter, with a weighting function for each frequency which is the spectrum of the series to be filtered. The spectrum is not known in practice but it can be easily estimated.

What we do is to solve a problem similar to that in Christiano and Fitzgerald (2003), but considering the use of other variables that can help predict the signal of interest. Suppose we are interested in isolating the fluctuations corresponding to the interval of frequencies $]\omega_l, \omega_h[$ of the series $\{x_t\}_{t=1}^{T}$. Suppose we have n series of covariates $z_1, ..., z_n$. To obtain the minimum mean squared error estimate of the ideally filtered series at time t, we choose weights $\{\widehat{B}_j^{p,f}, \widehat{R}_{1,j}^{p,f}, ..., \widehat{R}_{n,j}^{p,f}\}_{j=-f,...,p}$ associated with the series of interest and the available covariates, that solve the following problem:

$$\underset{\{\widehat{B}_j^{p,f},\widehat{R}_{1,j}^{p,f},...,\widehat{R}_{n,j}^{p,f}\}_{j=-f,...,p}}{Min} E[(y_t-\widehat{y_t})^2|\widehat{y_t}\in \pounds(x,z_1,...,z_n)] \tag{4}$$

where $\pounds(v)$ denotes the linear span generated by the vector random variable v, $x=(x_1,x_2,...,x_T)^{'}$, $z_1=(z_{1,1},z_{1,2},...,z_{1,T})^{'}$, ..., $z_n=(z_{n,1},z_{n,2},...,z_{n,T})^{'}$ and $y_t=B(L)x_t$ is the ideally filtered observation at time t. The estimate $\widehat{y_t}$ of the ideally filtered observation y_t is a weighted sum of past and future values of x_t and the covariates $z_1,...,z_n$:

$$\widehat{y_t}=\sum_{j=-f}^{p}\widehat{B}_j^{p,f}x_{t-j}+\sum_{s=1}^{n}\sum_{j=-f}^{p}\widehat{R}_{s,j}^{p,f}z_{s,t-j} \tag{5}$$

p denotes the number of observations in the past that are considered and f the number of observations in the future that are considered. Although we will present the solution for general values of p and f, the filter that uses all the observations in the various series will have $p=t-1$ and $f=T-t$ where T is the sample size. In the case of the CF filter we have only $\widehat{y_t}=\sum_{j=-f}^{p}\widehat{B}_j^{p,f}x_{t-j}$. In the remainder of the paper we will drop the superscript p,f for notational convenience. The problem can be conveniently formulated in the frequency domain. Define the following polynomials in z:

$$\widehat{B}(z)=\sum_{j=-f}^{p}\widehat{B}_jz^j,\quad \widehat{R}_s(z)=\sum_{j=-f}^{p}\widehat{R}_{s,j}z^j,\quad s=1,...,n \tag{6}$$

Using versions of (1) and (2) it is easy to verify that solving (4) is equivalent to solving the following problem:

$$\underset{\{\widehat{B}_j,\widehat{R}_{1,j},...,\widehat{R}_{n,j}\}_{j=-f,...,p}}{Min}\int_{-\pi}^{\pi}[B(e^{-i\omega})-\widehat{B}(e^{-i\omega}),-\widehat{R}_1(e^{-i\omega}),...,-\widehat{R}_n(e^{-i\omega})]$$

$$S_{x,z_1,...,z_n}(\omega)[B(e^{i\omega})-\widehat{B}(e^{i\omega}),-\widehat{R}_1(e^{i\omega}),...,-\widehat{R}_n(e^{i\omega})]^{'}d\omega \tag{7}$$

where $S_{x,z_1,...,z_n}(\omega) = \begin{bmatrix} S_x & S_{x,z_1} & ... & S_{x,z_n} \\ S_{z_1,x} & S_{z_1} & ... & S_{z_1,z_n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{z_n,x} & S_{z_n,z_1} & ... & S_{z_n} \end{bmatrix}$ denotes the spectral (or pseudo-spectral) matrix of the vector $(x_t, z_{1,t}, ..., z_{n,t})$. We make the following assumptions ir order to solve the problem in (7).

Assumption 1. *The vector* $(\Delta x_t, z_{1,t}, ..., z_{n,t})$*, where* $\Delta = 1 - L$*, is covariance-stationary.*

So, we do not deal with more than one unit root in x_t. Also, the pseudo-spectrum of x_t, S_x, has a pole only at zero frequency, we therefore abstract from poles at frequencies other than zero due to, e.g., nonstationary seasonal components. A more general specification can be envisaged, leading to the simultaneous solution of seasonal adjustment and signal extraction problems, but that is beyond the scope of this paper. The implied assumption that the covariates $z_{1,t}, ..., z_{n,t}$ are covariance-stationary is not restrictive, as long as we assume that no cointegration relations exist within the vector $(x_t, z_{1,t}, ..., z_{n,t})$, where now $z_{1,t}, ..., z_{n,t}$ are allowed to be integrated. Assuming this, suppose we have an indicator $z_{l,t}$ which is integrated of order 1. In this situation the polynomial $\widehat{R}_l(L)$ would need to have a unit root so that the term $\sum_{j=-f}^{p} \widehat{R}_{l,j}^{p,f} z_{l,t-j}$ from the solution to (7) is reduced to stationarity. Otherwise the criterion is infinite since some elements of $S_{x,z_1,...,z_n}(\omega)$ have a pole at zero frequency. This is also true for $\widehat{B}(L)$, that is, $\widehat{B}(L)$ must have a unit root since $B(1) = 0$. But this is equivalent to take initially first differences to the integrated series[2]. The solution depends only on the second order moments of $(\Delta z_l, \Delta x_t)'$, i.e., the information from the level of the integrated $z_{l,t}$ is irrelevant. In the presence of cointegration relations we would have stationary linear combinations of $x_t, z_{1,t}, ..., z_{n,t}$. This would allow us to have a finite variance solution $\widehat{y}_t = \sum_{j=-f}^{p} \widehat{B}_j^{p,f} x_{t-j} + \sum_{s=1}^{n} \sum_{j=-f}^{p} \widehat{R}_{s,j}^{p,f} z_{s,t-j}$, even in the presence of integrated $z_{l,t}$'s, without resorting to the unit-root restrictions $\widehat{B}(1) = \widehat{R}_l(1) = 0$. In principle, one could exploit such information and incorporate it in the solution. However, in practice we regard $z_{1,t}, ..., z_{n,t}$ as available indicators without much of a structural content.

[2]Apart from the sordid detail of loosing one observation in the beginning of the sample due to differencing.

Assumption 2. *The vector* $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ *follows an MA(M) process:*

$$\begin{bmatrix} \Delta x_t \\ z_{1,t} \\ \vdots \\ z_{n,t} \end{bmatrix} = \sum_{j=0}^{M} \Psi_j \varepsilon_{t-j},$$

where $\Psi_0 = I_{(n+1)\times(n+1)}$ *and* $\{\varepsilon_t\}$ *is a vector white noise sequence*[3].

The derivation of the solution in Appendix A can easily be adapted to consider the case $M = \infty$ along with $\sum_{j=0}^{\infty} |\Psi_j| < \infty$, but in practice we will only need the solution for finite M. This is because we will estimate the spectrum (or the autocovariance function) of $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ non-parametrically. Only a finite number (say M) of autocovariances can be used, although M is allowed to grow with T. If we estimated the process parametrically with, e.g., a VAR model, then we would have to determine the estimated Ψ_j, derive the autocovariance function and use the formulae provided here with $M = \infty$. Only the determination of the infinite sums in (a.4) (in Appendix A) could be cumbersome but easily approximated numerically. Note also that $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ has zero mean. In practice, this requires that $\Delta x_t, z_{1,t}, ..., z_{n,t}$ are normalized to have zero mean. In the case of x_t this is equivalent to initially removing a linear trend.

The solution to (7) under assumptions 1 and 2 is derived in Appendix A. For each observation in the sample, the weights of the filter are obtained by simply solving a linear system with $(p + f + 1) \times (n + 1)$ equations and unknowns. The solution depends on the second moments of $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ and on the weights of the ideal filter. Define $\widehat{B} = (\widehat{B}_p, \widehat{B}_{p-1}, ..., \widehat{B}_0, ..., \widehat{B}_{-f+1}, \widehat{B}_{-f})'$ and $\widehat{R}_s = (\widehat{R}_{s,p}, \widehat{R}_{s,p-1}, ..., \widehat{R}_{s,0}, ..., \widehat{R}_{s,-f+1}, \widehat{R}_{s,-f})'$, $s = 1, ..., n$. Stack these vectors in the vector of weights $\widehat{W} = (\widehat{B}', \widehat{R}_1', ..., \widehat{R}_n')'$. The linear system solved to recover the solution $\widehat{W}$ is the following:

$$V = Q\widehat{W} \tag{8}$$

[3] This assumption obviously implies assumption 1. We include it for ease of exposition.

where

$$V = \begin{bmatrix} S_{-f} & \cdots & S_{p-1} & \tilde{S}_p & 0 \end{bmatrix}'$$

and

$$Q = \begin{bmatrix} Q_{-f} \\ \vdots \\ Q_{p-1} \\ \tilde{Q}_p \\ U \end{bmatrix},$$

where

$$U = \begin{bmatrix} \underbrace{1 \;\; 1 \cdots 1}_{p+f+1} & \underbrace{0 \;\; 0 \cdots 0 \;\; 0 \;\; 0 \cdots 0}_{(p+f+1)\times n} \end{bmatrix}$$

The vectors $S_{-f}, ..., S_{p-1}$ are defined in (a.5) and (a.6) in Appendix A and the matrices $Q_{-f}, ..., Q_{p-1}$ are defined implicitly in (a.8). $\tilde{S}_p$ is just S_p as defined in (a.5) with the first element deleted and $\tilde{Q}_p$ is defined in the same way as the Q_j in (a.8) but with the first row deleted.

The case when all the points in the series are used amounts to having p and f varying with t. Specifically, $p = t - 1$ and $f = T - t$. Therefore, to get the weights that will be used to filter $\{x_t\}_{t=1}^{T}$ we need to solve the system in (8) T times. It is however easy to see that in this case the Q matrix in (8) will always be the same, it does not vary with t. Only the V vector varies with t.

The case when x_t is stationary reduces to a straightforward adaptation of the solution presented above, and it is also described in Appendix A. An algorithm in pseudo-code that constructs all the objects needed to solve the problem is presented in Appendix B.[4]

Remark 1. It is important to notice at this point that nothing in our solution is dependent on the specific ideal filter weights. Although we have in mind a specific signal extraction problem, the weights in (3) could be substituted by the weights of any (absolutely summable and stationary) distributed lag (or linear filter). This includes seasonal adjustment filters and the HP (infinite

[4]Mathematica code is available from the author upon request.

sample) filter. This follows from the fact that the derived solution to (4) does not rely on the weights of the (symmetric) ideal band-pass filter. We have further ensured that the solution is robust to asymmetry in the filter weights, although this may be of little practical interest.

Remark 2. If $B(1) \neq 0$, x_t has a unit root but $x_t - B(L)x_t$ is stationary, redefine the problem so as to isolate this stationary component. The redefined filter $B^*(L) = 1 - B(L)$ will obviously have a unit root ($B^*(1) = 0$). E.g., if one is interested in simply forecasting x_{T+1} (i.e. $B(L) = L^{-1}$), redefine the problem so as to forecast $x_{T+1} - x_T$.

Remark 3. Within this framework, as opposed to the state space approach of Valle e Azevedo, Koopman and Rua (2006), it is not possible to incorporate time series recorded at mixed frequencies, e.g., having quarterly x_t (say GDP) and monthly $z_{l,t}$'s (say Industrial production or Consumer confidence). Also, we do not deal with missing observations in the $z_{l,t}$'s, except if these are consecutive and in the end of the sample. In this case one should just trivially relabel the time subscript t, shifting the series so that they match the end of the sample for x_t.

Remark 4. It is definitely possible to extract the signal $y_{T+k} = B(L)x_{T+k}$ for $k \geq 0$. One just needs to set $f = -k$ in the solution, so that only the available information (that is, up to period T) is taken into consideration. Also, if you think the Bureau of statistics will revise x_T or other earlier estimates of the series of interest you can neglect them by choosing the appropriate f and relabelling the time subscript t for the $z_{l,t}$'s so that all the sample points of these series are considered. In this case, it might be reasonable to include the series of the first estimates of the Bureau of statistics as a $z_{l,t}$. Hopefully they are informative.

2.3 Other filters as particular cases of the derived filter

When $n = 0$, $p = f$ and constant for all t, $S_x(\omega)$ is constant for all ω and the restriction $\widehat{B}(1) = 0$ is also imposed we get the band-pass filter proposed by Baxter and King (1999). In this filter, information from other variables is not exploited, second order properties of the series are not exploited and $p = f$ observations are lost in the beginning and end of the sample. Since the ideal filter is independent of the particular representation of the time series, if $p = f$ is large the approximation is accurate. There is also an advantage in that the resulting filtered series is stationary, which may be important from an econometric point of view. If the purpose of the analyst is to look at historical or simulated data, without worrying about the estimates of the ideal filtered series near the end of the sample, using the BK filters is a very good option. Real-time estimates cannot however be computed.

When $n = 0$, we obtain exactly the band-pass filter of Christiano and Fitzgerald (2003). Our contribution is just extending that solution by exploiting information from an arbitrary number of covariates. It is clear from (4) that our solution always improves upon the Christiano and Fitzgerald (2003) solution if the true second order properties of the time series are known when applying both filters. If the second order properties are not known and need to be estimated, it is not clear that using covariates will always lead to more accurate estimates. Our conjecture is that if the covariates are highly correlated and lead dynamically the fluctuations of x_t in the specified range of frequencies, then improvements can be expected in the end of the sample, the observations of interest for the policy-maker. We will verify this conjecture in sections 3 and 4.

2.4 Previous time domain solutions

There is an interesting connection between the solution to (4) that Christiano and Fitzgerald (2003) derived for the univariate case and some earlier literature that used a time domain approach. Specifically, in a seasonal adjustment context, Geweke (1978) and Pierce (1980) present the time domain solution to the same univariate problem analyzed in Christiano and Fitzgerald (2003). It is shown that the best approximation to the ideal filter (in the case of Geweke 1978, and Pierce 1980, an arbitrary moving-average filter, as ours can be interpreted) is equivalent to applying the ideal filter to the series of interest, but with the particularity that this series is extended with optimal backcasts and forecasts (based on the available observations) when data points are not available. This extension of the finite sample is infeasible when the optimal filter to be approximated has infinite leads and lags, as is the case with an ideal band-pass filter. So, an advantage of the frequency domain approach by Christiano and Fitzgerald (2003) is that it provides a closed form solution in this situation.

It is important to observe that the existing literature to address the problem in (4) has only solved the problem for the univariate case. This is true for both the literature using a time domain approach as well as that using a frequency domain approach. So, the main contribution of this paper is the extension to the multivariate case. Our solution allows the use of an arbitrary number of covariates in the approximation to the desired signal.

2.5 Estimation of the spectrum

As referred already, in practice we do not know the second order properties of the variables being used. The autocovariances (or the spectrum) of $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ need to be estimated. We propose estimating the spectrum of $(\Delta x_t, z_{1,t}, ..., z_{n,t})'$ non-parametrically. The estimator is the following:

$$\widehat{S}_{\Delta x, z_1, ..., z_n}(\omega) = \frac{1}{2\pi}(\widehat{\Gamma}(0) + \sum_{k=1}^{M(T)} \kappa(k)(\widehat{\Gamma}(k)e^{i\omega k} + \widehat{\Gamma}(k)'e^{-i\omega k}))$$

where $\kappa(k)$ is a lag window, $M = M(T)$ is a function of the sample size T, and

$$\widehat{\Gamma}(k) = \frac{1}{T}\sum_{t=k+1}^{T}((\Delta x_t, z_{1,t}, ..., z_{n,t})')((\Delta x_{t-k}, z_{1,t-k}, ..., z_{n,t-k})')'$$

is the estimated autocovariance at lag k. $M(T) < T$ is denoted the truncation point and should grow at a rate slower than T if $\widehat{S}_{\Delta x, z_1, ..., z_n}(\omega)$ is to be consistent, that is, $M(T)/T \longrightarrow 0$ as $T \longrightarrow \infty$. Various lag windows have been proposed. In all the empirical applications we will use the Bartlett lag window for which $\kappa(k) = (1 - \frac{k}{M(T)+1})$. Given this estimator, we can use the formulae derived in Appendix A to obtain a filtered series. We just need to put the truncation point M as the M (recall, the moving average order) in those formulae and we also need to adjust all the (estimated) covariances, multiplying them by the factor $\kappa(k)$. Note that the mean was not subtracted in the estimation of the autocovariance function. This is because it was assumed that the mean was initially removed from the raw data. In the case of x_t, this is equivalent to initially removing a linear trend. Thus, in addition to the noise induced by the estimation of second moments, we have the noise in the estimation of the various means and of the linear trend of x_t. In the case of the HP filter or the BK filter removing the linear trend is not needed since they remove linear trends and also more than one unit root (precisely 2 unit roots in the case of the BK filter and 4 in the case of the HP filter). That property is lost in the CF filters and in the filters developed here if symmetry is lost, i.e., if $p \neq f$ which is the case if we want real-time estimates of the ideally filtered series. Thus, the requirement that $\widehat{B}(1) = 0$ only guarantees that one unit root is removed and a linear trend is transformed into a (generally) non-zero drift. The initial estimation of the linear trend corresponds to this drift adjustment.

3 Behavior of the filter and comparison with the univariate filter

We analyze now the behavior of the proposed multivariate band-pass filter in various contexts. It is clear that adding covariates provides more accurate estimates of the ideally band-pass filtered series, if the true second moment properties of the data are known. This is obvious from (4), the criterion can only decrease if we add covariates. Corner solutions (in which the weights assigned to a particular covariate would all be 0) would only arise if the series of interest and the covariate were uncorrelated or if one series was a linear combination of the other series. However, since second order moments need to be estimated in practice, it is not clear that always adding covariates or adding an arbitrary number of covariates results in a better signal. Arguably, if the covariates are highly correlated with x_t , the estimation of the spectrum is more precise and we should expect improvements. Also, if this relation is such that some or all covariates have leading properties w.r.t. x_t , then we should also expect improvements in the end of the sample. The problem of any band-pass filter in the end of the sample results from the ignorance about future events. The extrapolation of the future dynamics of x_t using covariates with leading properties in the type of fluctuations that we want to isolate should help improving the extracted signal. But again, it is not clear that the noise added by the estimation of the spectrum will not cancel those improvements. A simulation exercise was performed to try to answer these questions, providing guidelines about the general conditions that have to be met to usefully apply the filter. An analytical proof of the improvements, mixing the uncertainty of estimation with conditions on the dynamic relations between the covariates and x_t would be an enormous task, and is beyond the scope of this paper.

We will proceed as follows: First, we will compare the properties of filtered data assuming that we know some DGP's. This will give us an upper bound on the improvements that we can expect once we have to estimate second order moments. Second, in order to analyze the effects of estimating second moments, many realizations of those stochastic processes will be generated (specifically, 5000 realizations). An equivalent of our variable of interest, x_t, will be generated as well as covariates with varying stochastic properties. We will then apply an (almost) ideal filter to x_t. This requires dropping a considerable amount of observations in the endpoints of the simulated sample. We will then apply various filters, including the one developed here with estimated second moments, to the series that has the endpoints observations dropped. This mimics the unavailability of future data. Then we will compare in various observations near the end of this shorter sample the cross-sectional (across realizations of the stochastic processes) correlation with the ideally filtered series, the variance of the filtered series to analyze the extent of non-stationarity of these and also look at the phase effect between the ideally filtered series and the other filtered series.

We believe that the comparison with filters previously proposed in the literature is quite consolidated in Christiano and Fitzgerald (2003). Since the CF filter is the best performing band-pass filter as of now, we will focus on the comparison between the multivariate filter developed here and the CF filter.

We consider four variations of this basic setup. First, we will consider covariates without leading properties. Does this lead to any advantage if estimation noise is added? Then we will repeat the exercise with covariates showing leading properties. Also, a higher and reasonable number of covariates will be added[5]. Another variation deals with the strength of the covariation with the series of interest. In all these cases the sample size of the simulated data will be either $T = 200$ (say, 50 years of post-war quarterly data) or $T = 50$ (thinking in 50 years of annual data). For the computation of the almost ideally filtered series 200 additional data points will be generated in the case of the simulated quarterly data and 50 additional data points for the case of annual data. This almost ideally filtered series is the result of applying the BK filter with $p = f = 200$ for the quarterly case and $p = f = 50$ for the annual case. Two bands of periodicities were analyzed: The $[6, 32]$ periods band and the $[2, 8]$ periods band, thinking in quarterly and annual data. The approximation of these almost ideal filters to the ideal filters isolating the referred bands is very accurate (see figure 1).

3.1 Criteria for the evaluation and comparison of the filters

We look at the correlation between the ideally filtered series (y_t) and the various filtered series (denoted by $\widehat{y_t}$). In the case when the true second order moments are known we can see that, since $E[y_t] = E[\widehat{y_t}] = 0$:

$$E[(y_t - \widehat{y_t})^2 | \widehat{y_t} \in \pounds(x, z_1, ..., z_n)] \; = Var[y_t - \widehat{y_t}] = Var[y_t] + Var_t[\widehat{y_t}] - 2Cov_t[y_t, \widehat{y_t}]$$

Now, since $\widehat{y_t}$ solves a projection problem we can write $y_t = \widehat{y_t} + \varepsilon_t$, where ε_t is orthogonal to the elements in $(x, z_1, ..., z_n)$. Therefore, $Cov_t[y_t, \widehat{y_t}] = Var_t[\widehat{y_t}]$ and $E[(y_t - \widehat{y_t})^2 | \widehat{y_t} \in \pounds(x, z_1, ..., z_n)] = Var[y_t] - Var_t[\widehat{y_t}] = (1 - Corr_t[y_t, \widehat{y_t}]^2)Var[y_t]$ by straightforward arrangements. $Corr_t[y_t, \widehat{y_t}]$ is therefore a good measure of the variance of the approximation error. Also, given the expression for $Cov_t[y_t, \widehat{y_t}]$ it is easy to conclude that:

$$Corr_t[y_t, \widehat{y_t}] = \left[\frac{Var_t[\widehat{y_t}]}{Var[y_t]} \right]^{\frac{1}{2}}$$

[5] We have not considered more than 10 indicators because it becomes computationally difficult to assess the performance of the filter when data is simulated.

Whether or not indicators are used, the expression for $Var[y_t]$ can be easily calculated from the pseudo-spectrum of x_t. Using (1) and (2) we have:

$$Var[y_t] = 2\int_{\omega_l}^{\omega_h} S_x(\omega)d\omega$$

where $S_x(\omega)$ is the pseudo-spectrum of x_t. Also by (2) $Var_t[\widehat{y}_t]$ is given by:

$$Var_t[\widehat{y}_t] = [\widehat{B}(e^{-i\omega}), \widehat{R}_1(e^{-i\omega}), ..., \widehat{R}_n(e^{-i\omega})]S_{x,z_1,...,z_n}(\omega)[\widehat{B}(e^{i\omega}), \widehat{R}_1(e^{i\omega}), ..., \widehat{R}_n(e^{i\omega})]'$$

where $S_{x,z_1,...,z_n}(\omega)$ denotes the pseudo-spectral matrix of the vector $(x_t, z_{1,t}, ..., z_{n,t})$ and $\widehat{B}(e^{-i\omega})$, $\widehat{R}_1(e^{-i\omega})$,..., $\widehat{R}_n(e^{-i\omega})$ denote the Fourier transform of the polynomials in (6), with the weights given by the solution to the problem of determining $\widehat{y}_t$ in the various cases. In all the univariate filters the expression is modified to:

$$Var_t[\widehat{y}_t] = \widehat{B}(e^{-i\omega})S_x(\omega)\widehat{B}(e^{i\omega})$$

where $S_x(\omega)$ denotes the pseudo-spectrum of x_t. Therefore, if the second order properties of the data are known, we can calculate $Corr_t[y_t, \widehat{y}_t]$ and $Var_t[\widehat{y}_t]/Var[y_t]$ straightforwardly. We will also look at $Corr_t[y_{t-k}, \widehat{y}_t]$ in order to further assess the degree of non-stationarity of the filtered data. To get a closed form expression for $Corr_t[y_{t-k}, \widehat{y}_t]$, first form the vector $(x_t, z_{1,t}, ..., z_{n,t})$. Getting the filtered vector $(\widehat{y}_t, y_t)$ results by applying the filter $H_t(L)$ to $(x_t, z_{1,t}, ..., z_{n,t})$, where:

$$H_t(L) = \begin{bmatrix} \widehat{B}(L) \;\; \widehat{R}_1(L)...\widehat{R}_n(L) \\ B(L) \;\; 0 \;\; 0 \; ...0 \end{bmatrix}$$

By (2), the spectral matrix of $(\widehat{y}_t, y_t)$, denoted by $S_{y_t,\widehat{y}_t}(\omega)$ is given by:

$$S_{\widehat{y}_t,y_t}(\omega) = H_t(e^{-i\omega})S_{x,z_1,...,z_n}(\omega)H_t'(e^{i\omega})$$

where $S_{x,z_1,...,z_n}$ is the pseudo-spectral matrix of $(x_t, z_{1,t}, ..., z_{n,t})$. Therefore, by (1), we have:

$$Cov_t[y_{t-k}, \widehat{y_t}] = \int_{-\pi}^{\pi} e^{i\omega k} S_{\widehat{y_t},y_t}(\omega) d\omega, k = 0, \pm 1, \pm 2, ...$$

which together with $Var[y_t]$ and $Var_t[\widehat{y_t}]$ can be used to straightforwardly compute $Corr_t[y_{t-k}, \widehat{y_t}]$. If $\widehat{y_t}$ is close enough to y_t then $Corr_t[y_{t-k}, \widehat{y_t}]$ should resemble the autocorrelation function of y_t, which is obviously symmetric around $k = 0$. The form of the asymmetry can be used to assess the phase effect between y_t and $\widehat{y_t}$. This phase effect is typical and obviously positive (y_t leads $\widehat{y_t}$) in the endpoints of the sample, due to the nature of the one-sided and asymmetric filter in those points.

Note again that the above expressions are used when we know the true second moments of the data. When these moments need to be estimated, we evaluate the behavior of the filters by using simulated data to compute these statistics.

3.2 Coincident Covariates

We analyze first coincident covariates by considering the following simple DGP:

$$\begin{bmatrix} \Delta x_t \\ z_{1,t} \\ \vdots \\ z_{n,t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{\Delta x,t} \\ \varepsilon_{z_1,t} \\ \vdots \\ \varepsilon_{z_n,t} \end{bmatrix} = \varepsilon_t \text{ where } \varepsilon_t \sim NID(0, \Sigma) \tag{9}$$

For the low correlation indicators all the diagonal elements of Σ are set equal to 1, the elements in the first row and first column equal to 0.5 and all the remaining elements equal to 0.4. For the highly correlated indicators, the elements in the first row and first column equal 0.95 and all the remaining elements equal 0.9. We have therefore x_t following a random walk, and n indicators only contemporaneously correlated with each other and with the innovation for Δx_t. Whenever second moments are estimated we set $M = 0$ (only the matrix Σ is estimated), thereby avoiding the estimation of the zero autocovariances of higher order (see also the discussion in section 3.4).

Designation	DGP	Σ Matrix	n
Coincident-Low Corr.-3 Indicators	in (9)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1, ..., n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1, ..., n$	3
Coincident-High Corr.-3 Indicators	in (9)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1, ..., n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1, ..., n$	3
Coincident-Low Corr.-10 Indicators	in (9)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1, ..., n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1, ..., n$	10
Coincident-High Corr.-10 Indicators	in (9)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1, ..., n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1, ..., n$	10

Table 1: DGP's considered: Coincident Indicators Case

Filter Designation	Moments Estimated	Multivariate	M
CF	No	No	0
Multivariate	No	Yes	0
Multivariate Estimation	Yes	Yes	0

Table 2: Filters considered: Coincident Indicators Case

The variations considered in the exercise are summarized in table 1.

It is important to notice that the optimal CF filter in this case is the filter recommended by Christiano and Fitzgerald (2003) for unit-root processes, even if the details of the second order properties are not known precisely (CF assuming the DGP is simply a pure random walk). It is argued that only small gains are obtained by estimating these moments. We will again verify this fact and argue that under certain conditions this does not hold for our multivariate filter. The details of the filters applied (either in simulated data or theoretically) are in table 2.

3.2.1 Results: [2,8] periods band

In this case, the results are as follows: there isn't any noticeable advantage in using the indicators in comparison to the univariate CF filter, even if the matrix Σ is known. However, adding covariates and estimating Σ does not lead to any relevant deterioration of the signal extracted by the multivariate filter. This is true whether the correlation is high or low and also if more covariates are used. For all the filters, the variance stabilizes rapidly and the correlation with the ideally filtered series can only be problematic in the last observation. Also, there is only a relevant asymmetry of $Corr_t[y_{t-k}, \widehat{y}_t]$ in the last observation[6]. Since the results are almost

[6] In this case, and in the remainder of the analysis, we will only show $Corr_t[y_{t-k}, \widehat{y}_t]$ for the *Multivariate* filter case. The figures for the other filters are very similar, with the exception of the case $k = 0$, which is fortunately

exactly the same across the variations considered in table 1, we only present the results for the *Coincident-Low Corr.-3 Indicators* case (figure 2).

We have also considered coincident indicators with richer properties, meaning that these are not only contemporaneously correlated with Δx_t. Coincident indicators are defined here as indicators with a negligible or zero phase effect with respect to Δx_t. No relevant improvements were again detected in theory, or deterioration once moments had to be estimated.

3.2.2 Results: [6,32] periods band

In the $[6, 32]$ periods case there are again no relevant differences across the filters used. No relevant improvements are achieved theoretically (if second moments are known) and no relevant deterioration is verified if Σ has to be estimated. This is true across all the variations considered in table 1. In all the cases, the correlation with the ideally filtered series only drops below 0.8 in the last 6 observations of the sample and the variance also stabilizes quite rapidly, although it reaches only 50% of the variance of the ideally filtered series in the final point. Again, in figure 3 we present the various analyzed statistics, again only for the *Coincident-Low Corr.-3 Indicators* case. Notice also that $Corr_t[y_{t-k}, \widehat{y_t}]$ for the *Multivariate* filter is still asymmetric in the third last observation, although slightly, confirming the intuitive idea that bands with larger upper periods lead to a phase effect that is relevant in more of the final observations.

Again, the consideration of indicators with richer (coincident) dynamic properties did not lead to relevant improvements of the *Multivariate* filter. Also, estimation of second moments was not particularly harmful. This seems indeed to be a robust finding.

3.3 Leading Covariates

We define leading indicators as indicators with positive phase effect when compared to Δx_t. We consider the following process:

shown in a separate figure for all the filters and for more observations. This is done for reasons of parsimony of plots.

$$
\begin{bmatrix} \Delta x_t \\ z_{1,t} \\ \vdots \\ z_{n,t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{\Delta x,t} \\ \varepsilon_{z_1,t} \\ \vdots \\ \varepsilon_{z_n,t} \end{bmatrix} + \sum_{j=1}^{M} \Psi_j \begin{bmatrix} \varepsilon_{\Delta x,t-j} \\ \varepsilon_{z_1,t-j} \\ \vdots \\ \varepsilon_{z_n,t-j} \end{bmatrix}, \text{where } \begin{bmatrix} \varepsilon_{\Delta x,t} \\ \varepsilon_{z_1,t} \\ \vdots \\ \varepsilon_{z_n,t} \end{bmatrix} \sim NID(0, \Sigma) \qquad (10)
$$

where $M = 4$. We consider using 3 or 10 covariates with leading properties. We achieve this by parametrizing the DGP such that the cross-correlation function between Δx_t and each of the indicators has a maximum at a positive lag, that is, the indicators lead in fact Δx_t. For the case of annual data, we consider a lead reflected in a maximum occurring at lag 1 (denoted by $k = 1$) in the cross-correlation function of $(\Delta x_t, z_{i,t})$, $Corr[\Delta x_t, z_{i,t-k}]$. In the case of quarterly data we consider this maximum occurring at lag 3 ($k = 3$). Again, for the low correlation indicators all the diagonal elements of Σ are set equal to 1, the elements in the first row and first column equal to 0.5 and all the remaining elements equal to 0.4. For the highly correlated indicators the elements in the first row and first column equal to 0.95 and all the remaining elements equal to 0.9. In the case of annual data this implies that the maximum in the cross-correlation function of $(\Delta x_t, z_{i,t})$ (at lag 1) is either 0.6 or 0.85. In the case of quarterly data the value achieved at lag 3 (the maximum) is either 0.55 or 0.8. These are values that can realistically be achieved in practice. In the estimation, we set $M = 4$, meaning that we disregard possible misspecification in the lag length (see also discussion in section 3.4). The variations considered in this highly specific and only illustrative exercise are summarized in table 3.

It is very important to notice that the parametrization is such that the univariate representation of the variable of interest (Δx_t) is always the same for each band of interest. This is done for sake of comparability with the univariate filters considered and also to analyze the effect of the variations in the Σ matrix and in the number of indicators more precisely. Also, we have ensured that the cross-correlation function of $(\Delta x_t, z_{i,t}), i = 1, ..., n$ is almost unchanged once we add more covariates (n increases). That is, if n increases we can say that we have more indicators with similar "quality".

The details of the filters applied (either in simulated data or theoretically) are in table 4.

Designation	**Band**	**DGP and** $Corr[\Delta x_t, z_{i,t-k}]$	Σ **Matrix**	**n**
Leading-Low Corr.-3 Indicators	$[2,8]$	in(10) fig. 4 (i)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1,...,n$	3
Leading-High Corr.-3 Indicators	$[2,8]$	in(10) fig. 4 (ii)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1,...,n$	3
Leading-Low Corr.-10 Indicators	$[2,8]$	in(10) fig. 4 (iii)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1,...,n$	10
Leading-High Corr.-10 Indicators	$[2,8]$	in(10) fig. 4 (iv)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1,...,n$	10
Leading-Low Corr.-3 Indicators	$[6,32]$	in(10) fig. 5 (i)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1,...,n$	3
Leading-High Corr.-3 Indicators	$[6,32]$	in(10) fig. 5 (ii)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1,...,n$	3
Leading-Low Corr.-10 Indicators	$[6,32]$	in(10) fig. 5 (iii)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.5, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.4, i \neq j = 1,...,n$	10
Leading-High Corr.-10 Indicators	$[6,32]$	in(10) fig. 5 (iv)	Diagonal elements are 1 $Cov[\varepsilon_{\Delta x,t}, \varepsilon_{z_i,t}] = 0.95, i = 1,...,n$ $Cov[\varepsilon_{z_j,t}, \varepsilon_{z_i,t}] = 0.9, i \neq j = 1,...,n$	10

Table 3: DGP's considered: Leading Indicators Case

Filter Designation	**Moments Estimated**	**Multivariate**	**M**
CF	No	No	4
CF Estimation	Yes	No	4
CF Unit Root	No	No	0
Multivariate	No	Yes	4
Multivariate Estimation	Yes	Yes	4

Table 4: Filters considered: Leading Indicators Case

3.3.1 Results: [2,8] periods band

The statistics used to compare the performance of the various filters are presented in figures 6 and 7. To avoid a multiplication of plots, we report the results only for the worst and best case scenarios, which turned out to be the *Leading-Low Corr.-3 Indicators* case and the *Leading-High Corr.-10 Indicators* case respectively. In the comments below we make use of unreported results.

The first thing to notice is that the multivariate filter now definitely outperforms the univariate filters, theoretically and also when second moments need to be estimated. However, estimation makes the gains far from relevant when the covariates have only a moderate correlation with Δx_t. But when the indicators are highly correlated with Δx_t, the performance of *Multivariate Estimation* is very close to the performance of *CF* (in which case the second moments are known). Also, the addition of more indicators results in a better performance (theoretically and in practice), but the improvements are very slight. The relevant improvements arise when the indicators are strongly correlated with Δx_t, instead of moderately. Thus, in this example, the multivariate filter in the *Leading-High Corr.-3 Indicators* case beats the multivariate filter applied to the *Leading-Low Corr.-10 Indicators* case. The figures for these cases would lie "between" figures 6 and 7.

These results are very promising. With the use of reasonably good indicators the theoretical improvements are relevant and estimation still leads to improvements of the extracted signal. In any case, the general patterns are very similar to the ones obtained in section 3.1. Only the last observation could cause some concern, since the variance and the correlation with the ideally filtered series drop clearly in this observation. Also, there is only a relevant asymmetry of $Corr_t[y_{t-k}, \widehat{y_t}]$ in the last observation. *CF*, *CF Estimation*, *Multivariate* and *Multivariate Estimation* all beat clearly, in the last observation, *CF Unit Root*. Thereafter all filters perform similarly.

3.3.2 Results: [6,32] periods band

The statistics used to compare the performance of the various filter are presented in figures 8 and 9. Again, we report the results only for the worst and best case scenarios, which turned out to be (again) the *Leading-Low Corr.-3 Indicators* case and the *Leading-High Corr.-10 Indicators* case respectively.

In this case, the multivariate filter considerably outperforms the univariate filter, both theoretically and even when second moments need to be estimated. The performance of *Multivariate Estimation* is indeed much closer to the performance of *Multivariate*[7]. Moreover, *Multivariate Estimation* clearly beats *CF* (for which the true moments are actually known!). Also, *CF* and *CF Estimation* have now a performance close to *CF Unit Root*. The general patterns are very similar to the ones obtained in section 3.1. Only in the last 6 observations is the performance

[7] This should not be surprising since the sample size for this simulated "quarterly" data is $T = 200$ instead of $T = 50$ for the "annual" data considered before. This, together with the fact that the DGP´s have similar properties in both cases, makes estimation of second moments more precise.

of all the filters worrying, since the variance and the correlation with the ideally filtered series drops fast in these observations. The asymmetry of $Corr_t[y_{t-k}, \widehat{y_t}]$ in the three last observations is again noticeable in all the filters.

As in the case of annual data, the big improvements in the multivariate filter arise when the indicators are strongly correlated with Δx_t, instead of moderately. The addition of more indicators results again in a better performance (theoretically and in practice), but the improvements are slight. So, as in the case of annual data, the multivariate filter in the *Leading-High Corr.-3 Indicators* case beats the multivariate filter applied to the *Leading-Low Corr.-10 Indicators* case. The figures for these cases would lie "between" figures 8 and 9. The main fact is that *Multivariate Estimation* clearly beats the univariate filters, even when the true second moments are used in the later. The key seems therefore to have good (leading) indicators at hand. We will see in section 4 that these indicators are available in practice.

3.4 Note on the choice of M

We have not considered, in the exercises above, misspecification of the moving average process order (M). We repeated the exercise considering the estimation of $\sqrt{T}$ autocovariances (meaning that those greater than M were not fixed equal to 0, but estimated) and the conclusions remained unchanged. The estimation of all the zero autocovariances of higher order was only slightly harmful to both the multivariate filter and the CF filter.

In order to choose M, another option would be to sequentially test if the autocovariance matrices have all elements equal to 0 and set M as the order of the last "significant" autocovariance matrix. Better advice on how to estimate the spectrum (including the advantages, according to the situation, of various lag windows) can be found in Priestley (1981).

4 Application: A business cycle indicator

We use the developed filter to construct a business cycle indicator for the U.S. economy. The series of interest in this case is quarterly GDP, the best available proxy of aggregate economic activity.

The chosen covariates appear to have leading properties in the 6 to 32 quarters periodicities band (see Stock and Watson 1999). These covariates are the Help Wanted adds index, Industrial Production index, Capacity Utilization, Average weekly working hours, Non-Farm Output and Hours of All Persons (Business Sector) (see Appendix B for details). It should be noted that we have not provided a criterion to select a subset of available covariates. This selection should be done because it is not clear that including as many covariates as possible will result in a superior performance of the multivariate filter. In section 3 we have only shown that it is *possible* to improve the performance of the filter by including more covariates. In practice, given a set of covariates, we suggest a subset selection based on the real-time performance of the possible multivariate filters (using one or more of the criteria below).

In figure 10 we plot the decomposition of U.S. GDP into three components, using our multivariate band-pass filter. These are the trend, which corresponds to the band of periodicities (quarters in this case) $]32, \infty[$, the business cycle fluctuations (our indicator) which correspond to the band $[6, 32]$ time periods and the noise component containing the fluctuations with period less than 6 time periods[8].

The performance of the proposed business cycle indicator will be assessed by analyzing it's real-time performance. Specifically, and in line with Orphanides and Van Norden (2002), we will look at the revisions observed by using our method. In practice, once new data is available, the filtered estimates vary near the end of the sample. This variation is due to revisions in the data itself, which we do not analyze here, and revisions due to the nature of the one sided filter used in the end of the sample. The magnitude of the revisions if often large, whatever the method used, even in a multivariate context (Orphanides and Van Norden 2002). It is therefore crucial to assess the importance of these revisions when using the proposed filter. We will additionally compare our approach with the HP and CF univariate filters. In the case of our filter and the CF filter, we will eliminate another source of possible revision. The needed second order moments will be the ones obtained using the whole sample. It is expected that the variation stemming

[8]In practice, we apply the multivariate filter to isolate the $[2, 6]$ periods band and the $[6, 32]$ periods band. The trend is calculated as the original series minus these fluctuations.

from second moments uncertainty will be reduced as the sample size gets larger, i.e., from today onwards. We have however verified that the results are only slightly worse for both filters if we estimate the moments in real-time. More importantly, the ranking of the filters remains unchanged.

We will compare the final filtered series, the one that uses all the data available, with the real-time filtered series, the series obtained by using each of the filters conditional on knowing only the data available at each point in time. It is in fact confirmed that whichever method is used, the real-time assessment of the cyclical position is extremely hard. We plot in figure 11 the real-time and final estimates of the various business cycle indicators. In this analysis, and in view of the results obtained in the previous section, we do not consider the last 12 observations. We do this because the final filtered series would not be actually "final", but could vary a bit once new data became available. Also, 5 years of data are disregarded in the beginning of the sample. Again, the first filtered observations would vary if we had past data. Also, it would be unreasonable to use a very low number of data points in the computation of real-time estimates. However, these are obviously used in the computation of the final filtered series for the other observations.

We will compare the HP filter with smoothing parameter $\lambda = 1038$[9], the CF filter assuming that the true DGP is a random walk (*CF Unit Root*), the CF filter using the estimated second moments (*CF Estimation* with $M = 6$) and our multivariate filter (*Multivariate Estimation* with $M = 6$). The choice of $M = 6$ results from the inspection of the autocorrelation and cross-correlations of the various series. In addition to the plots, we assess the magnitude of the revisions by looking at the correlation between the final and the real-time estimates, the noise-to-signal ratio (the ratio of the standard deviation of the revisions to the standard deviation of the final estimate) and by the sign concordance (the proportion of time in which final and real-time estimates share the same sign). Table 5 presents the results. The HP filter is clearly

[9]The smoothing parameter is set to $\lambda = 1038$ instead of the typical $\lambda = 1600$ for quarterly data because this provides a closer approximation to the ideal filter isolating the $[2, 32]$ periods band when the time series process contains (almost) one unit root (see Pedersen 2001). Also, we are a bit unfair to the HP filter as we do not extend the series with backcasts and forecasts which would improve its performance (see Kaiser and Maravall 1999).

Method	Correlation		noise to signal		sign Concordance	
	72-02	88-02	72-02	88-02	72-02	88-02
HP filter	0.50	0.31	1.02	1.26	0.56	0.47
CF Unit Root	0.77	0.68	0.65	0.73	0.71	0.65
CF Est.	0.78	0.69	0.64	0.72	0.73	0.67
Multivariate Est.	0.82	0.75	0.59	0.67	0.84	0.84

Table 5: Correlation is the contemporaneous correlation between the real-time estimates and the final estimates of the business cycle. Noise-to-signal ratio is the ratio of the standard deviation of the revisions against the standard deviation of the final estimates. Sign concordance is the proportion of time in which final and real-time estimates share the same sign. The statistics are reported for the sample periods of 1972-2002 (72-02) and 1988-2002 (88-02).

the one that performs worse, being also very noisy (see figure 11). This was obviously expected since the HP filter does not eliminate high frequency fluctuations. *CF Estimation* improves only slightly upon *CF Unit Root* and our multivariate filter outperforms by relevant margin the best performing univariate filter. It seems that using leading indicators can be useful in determining more accurately the cyclical position. Also, there is room to consider even more and better indicators. Our search was not exhaustive and our objective is exemplificative.

Additionally, we look at the behavior of the estimates once $1, 2, ..., 6$ new observations are available. That is, we compare the estimates obtained with those additional data points with the final estimate. Once the new data is available, all the measures analyzed before should be more favorable for all the filters. In figure 12 we present the results. In the horizontal axis, 0 represents the real-time estimate (results already in table 5), 1 represents the estimate obtained when one future data point is available and so forth. Clearly, the multivariate band-pass filter outperforms all the other filters in every dimension. There are no relevant improvements if we estimate the process in the univariate case, *CF Unit Root* is already reasonable, but using information from the covariates leads to relevant improvements. Only in the case of the sign concordance is the multivariate filter beaten, but only after 4 and 6 new observations are available and the difference is negligible. After 6 new observations, all the filters perform similarly in the dimensions analyzed.

5 Concluding remarks

Accurate real-time estimates of business cycle fluctuations, or of any other signal that embodies unavailable information, can be extremely hard to get. Only with some projection of future developments can we assess a relative position with regard to the available past events and the indeterminate future events. It seems that any improvement in the methods used to determine this cyclical position must rely on multiple sources of information. We have followed this hint and confirmed that relevant improvements can arise. Our multivariate filter outperforms in various dimensions the avalilable univariate band-pass filters, given that certain conditions on the relations between our series of interest and available covariates are met.

We concluded that the use of covariates with leading properties with respect to the series of interest allows us to better distinguish what is signal and what is noise or what is signal and what is a long-term movement. This holds even if we need to estimate the second order moments of the data. It was not confirmed that in general more estimation (i.e., more covariates) deteriorates the extracted signal, on the contrary. But there is eventually a trade-off between more (imprecise) estimation on one hand and more information on the other. So, the question that remains to be answered is to what degree and under what conditions can the ignorance about certain properties of the data undermine the otherwise trivial result that states that more (non-redundant) information leads always to a more accurate signal extraction. Our simulation exercise gave us only a hint about the properties of the data that lead to (relevant) improvements.

Appendix A: Derivation of the filter

We want to solve the problem in (7)under assumptions 1 and 2. Since x_t has a unit root, all the elements in the first row and in the first column of the pseudo-spectral matrix, $S_{x,z_1,\ldots,z_n}(\omega)$, have a pole at zero frequency. Since $B(1) = 0$, this implies that $\widehat{B}(1) = 0$ (otherwise the criterion is infinite). This gives us only 1 equation in the $(n + 1) \times (p + f + 1)$ unknowns. Given this condition, the polynomial $\widehat{B}(z)$ can be written as follows:

$$\widehat{B}(z) = (1-z)b(z)$$

where $b(z) = b_{p-1}z^{p-1} + b_{p-2}z^{p-2} + ... + b_0 + ... + b_{-f+1}z^{-f+1} + b_{-f}z^{-f}$ with:

$$b_j = -\sum_{i=j+1}^{p} \widehat{B}_i \quad j = -f, ..., p-1$$

Now define

$$\widehat{B} = (\widehat{B}_p, \widehat{B}_{p-1}, ..., \widehat{B}_0, ..., \widehat{B}_{-f+1}, \widehat{B}_{-f})'$$

$$\widehat{R}_s = (\widehat{R}_{s,p}, \widehat{R}_{s,p-1}, ..., \widehat{R}_{s,0}, ..., \widehat{R}_{s,-f+1}, \widehat{R}_{s,-f})', s = 1, ..., n$$

$$b = (b_{p-1}, b_{p-2}, ..., b_0, ..., b_{-f+1}, b_{-f})' \tag{a.1}$$

The relation between $\widehat{B}$ and b can be represented as follows:

$$D\widehat{B} = b \tag{a.2}$$

where

$$D = \begin{bmatrix} -1 & 0 & 0 & 0 & ... & 0 & 0 \\ -1 & -1 & 0 & 0 & ... & 0 & 0 \\ -1 & -1 & -1 & 0 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & ... & -1 & 0 \end{bmatrix}_{(p+f)\times(p+f+1)}$$

This will be convenient when representing the solution to the problem as the solution to a linear system. Define:

$$\alpha(\omega) = [\overline{B}(e^{-i\omega}) - b(e^{-i\omega}), -\widehat{R}_1(e^{-i\omega}), ..., -\widehat{R}_n(e^{-i\omega})], \text{where } \overline{B}(z) = B(z)/(1-z)$$

recalling that:

$$\widehat{B}(z)=\sum_{j=-f}^{p}\widehat{B}_j z^j,\quad \widehat{R}_s(z)=\sum_{j=-f}^{p}\widehat{R}_{s,j}z^j,\quad s=1,...,n$$

The problem can be rewritten as follows:

$$\underset{\{b_j,\widehat{R}_{1,j},...,\widehat{R}_{n,j}\}_{j=-f,...,p-1},\{R_{1,p},...,R_{n,p}\}}{Min}\int_{-\pi}^{\pi}\alpha(\omega)S_{\Delta x,z_1,...,z_n}(\omega)\alpha(-\omega)'d\omega$$

where:

$$S_{\Delta x,z_1,...,z_n}(\omega)=\begin{bmatrix}(1-e^{-i\omega}) & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & 1 & 0\\ 0 & 0 & \cdots & 0 & 1\end{bmatrix}S_{x,z_1,...,z_n}(\omega)\begin{bmatrix}(1-e^{i\omega}) & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & 1 & 0\\ 0 & 0 & \cdots & 0 & 1\end{bmatrix}$$

represents the (now well defined) spectral matrix of the vector $(\Delta x,z_1,...,z_n)$. Define the vector $W_j=(b_j,\widehat{R}_{1,j},...,\widehat{R}_{n,j})'$, $j=-f,...,p-1$. It is straightforward to verify that:

$$\frac{\partial\alpha(\omega)}{\partial W_j}=-e^{-i\omega j}I_{(n+1)\times(n+1)}\text{ and }\frac{\partial\alpha(-\omega)'}{\partial W_j}=-e^{i\omega j}I_{(n+1)\times(n+1)}$$

Therefore, the first order conditions w.r.t. W_j are the following:

$$\int_{-\pi}^{\pi}[\frac{\partial\alpha(\omega)}{\partial W_j}S_{\Delta x,z_1,...,z_n}(\omega)\alpha(-\omega)'+\alpha(\omega)S_{\Delta x,z_1,...,z_n}(\omega)\frac{\partial\alpha(-\omega)'}{\partial W_j}]d\omega=0,$$

$$j=-f,-f+1,...,p-1$$

$$\Leftrightarrow\int_{-\pi}^{\pi}[e^{-i\omega j}S_{\Delta x,z_1,...,z_n}(\omega)\begin{bmatrix}\overline{B}(e^{i\omega})\\ 0\\ \vdots\\ 0\end{bmatrix}+$$

$$e^{i\omega j}\begin{bmatrix}\overline{B}(e^{-i\omega}) & 0 & \dots & 0\end{bmatrix} S_{\Delta x,z_1,\dots,z_n}(\omega)]d\omega =$$

$$\int_{-\pi}^{\pi}[e^{-i\omega j} S_{\Delta x,z_1,\dots,z_n}(\omega)\begin{bmatrix} b(e^{i\omega}) \\ \widehat{R}_1(e^{i\omega}) \\ \vdots \\ \widehat{R}_n(e^{i\omega})\end{bmatrix} +$$

$$e^{i\omega j}\begin{bmatrix} b(e^{-i\omega}) & \widehat{R}_1(e^{-i\omega}) & \dots & \widehat{R}_n(e^{-i\omega})\end{bmatrix} S_{\Delta x,z_1,\dots,z_n}(\omega)]d\omega,$$

$$j = -f, -f+1, ..., p-1$$

The second term inside the integrals in both sides of the above equations is just the complex conjugate of the first term. To conclude this, we use the fact that the spectral matrix $S_{\Delta x,z_1,\dots,z_n}(\omega)$ is Hermitian (equal to its conjugate transposed). The integral of the first term is therefore equal to the integral of the second term which implies that the equations can be reduced to:

$$\int_{-\pi}^{\pi}[e^{-i\omega j}\begin{bmatrix} S_{\Delta x}(\omega) \\ S_{z_1,\Delta x}(\omega) \\ \vdots \\ S_{z_n,\Delta x}(\omega)\end{bmatrix}\overline{B}(e^{i\omega})]d\omega =$$

$$\int_{-\pi}^{\pi}[e^{-i\omega j} S_{\Delta x,z_1,\dots,z_n}(\omega)\begin{bmatrix} b(e^{i\omega}) \\ \widehat{R}_1(e^{i\omega}) \\ \vdots \\ \widehat{R}_n(e^{i\omega})\end{bmatrix}]d\omega,$$

$$j = -f, -f+1, ..., p-1 \tag{a.3}$$

These are $(p+f)\times(n+1)$ equations in the $(p+1+f)\times(n+1)$ unknowns. Put:

$$\overline{S}_{\Delta x,z_1,\dots,z_n}(\omega) = (S_{\Delta x}(\omega), S_{z_1,\Delta x}(\omega), ..., S_{z_n,\Delta x}(\omega))'$$

Denote the left hand side of (a.3) by S_j . Subtract S_j to S_{j-1} to get:

$$\int_{-\pi}^{\pi}[(e^{-i\omega j}-e^{-i\omega(j-1)})\overline{S}_{\Delta x,z_1,\dots,z_n}(\omega)\overline{B}(e^{i\omega})]d\omega=$$

$$\int_{-\pi}^{\pi}[e^{-i\omega j}\overline{S}_{\Delta x,z_1,\dots,z_n}(\omega)(1-e^{i\omega})\overline{B}(e^{i\omega})]d\omega=$$

$$\int_{-\pi}^{\pi}[e^{-i\omega j}\overline{S}_{\Delta x,z_1,\dots,z_n}(\omega)B(e^{i\omega})]d\omega=V_j \tag{a.4}$$

$$j=-f+1,\dots,p-1$$

In order to get, as much as possible, closed form expressions for all the objects involved in the solution of the problem, we recall that:

$$\begin{aligned}\int_{-\pi}^{\pi}e^{i\omega j}d\omega &= 0, j=\pm1,\pm2,\dots\\ &= 2\pi, j=0\end{aligned}$$

Therefore (a.4) is just 2π times the coefficient on z^0 of:

$$B(z)z^{-j}\overline{S}_{\Delta x,z_1,\dots,z_n}(z)$$

where $\overline{S}_{\Delta x,z_1,\dots,z_n}(z)$ denotes a column of the autocovariance generating function. Now,

$$\overline{S}_{\Delta x,z_1,\dots,z_n}(z)=\begin{bmatrix}S_{\Delta x}(z)\\ S_{z_1,\Delta x}(z)\\ \vdots\\ S_{z_n,\Delta x}(z)\end{bmatrix}=\frac{1}{2\pi}.\begin{bmatrix}\gamma_{\Delta x}(0)+\gamma_{\Delta x}(1)(z+z^{-1})+\dots+\gamma_{\Delta x}(M)(z^M+z^{-M})\\ \gamma_{z_1,\Delta x}(-M)z^{-M}+\dots+\gamma_{z_1,\Delta x}(0)+\dots+\gamma_{z_1,\Delta x}(M)z^M\\ \vdots\\ \gamma_{z_n,\Delta x}(-M)z^{-M}+\dots+\gamma_{z_n,\Delta x}(0)+\dots+\gamma_{z_n,\Delta x}(M)z^M\end{bmatrix}$$

where $\gamma_{a,b}(j)$ is the autocovariance between a and b at lag j. Therefore, (a.4) is:

$$\begin{bmatrix} B_j\gamma_{\Delta x}(0)+\sum\limits_{i=1}^{M}(B_{j+i}+B_{j-i})\gamma_{\Delta x}(i) \\ B_j\gamma_{z_1,\Delta x}(0)+\sum\limits_{i=1}^{M}(B_{j-i}\gamma_{z_1,\Delta x}(i)+B_{j+i}\gamma_{z_1,\Delta x}(-i)) \\ \vdots \\ B_j\gamma_{z_n,\Delta x}(0)+\sum\limits_{i=1}^{M}(B_{j-i}\gamma_{z_n,\Delta x}(i)+B_{j+i}\gamma_{z_n,\Delta x}(-i)) \end{bmatrix} = V_j,\ j=-f+1,...,p-1$$

which is robust to asymmetry in the ideal filter weights. The S_j can then be obtained recursively by:

$$S_j = S_{j-1} + V_j, j = -f+1, ..., p \tag{a.5}$$

The only term for which there is no closed form expression is S_{-f}, which can easily be obtained numerically. We have:

$$S_{-f} = \int\limits_{-\pi}^{\pi}[e^{i\omega f}\overline{S}_{\Delta x,z_1,...,z_n}(\omega)\frac{B(e^{i\omega})}{1-e^{i\omega}}]d\omega = \tag{a.6}$$

$$\int\limits_{-\omega_h}^{-\omega_l}[e^{i\omega f}\overline{S}_{\Delta x,z_1,...,z_n}(\omega)\frac{1}{1-e^{i\omega}}]d\omega+$$

$$\int\limits_{\omega_l}^{\omega_h}[e^{i\omega f}\overline{S}_{\Delta x,z_1,...,z_n}(\omega)\frac{1}{1-e^{i\omega}}]d\omega$$

where the last equality is only valid for ideal band-pass filters. Similarly, the right hand side of (a.3) is just 2π times coefficient on z^0 of:

$$z^{-j}S_{\Delta x,z_1,...,z_n}(z)\begin{bmatrix} b(z) \\ \widehat{R}_1(z) \\ \vdots \\ \widehat{R}_n(z) \end{bmatrix}$$

where $S_{\Delta x,z_1,...,z_n}(z)$ denotes the autocovariance generating function. The first element of this vector is:

$$z^{-j}S_{\Delta x}(z)b(z)+\sum_{s=1}^{n}z^{-j}S_{\Delta x,z_s}(z)\widehat{R}_s(z) \tag{a.7}$$

If $p-1-M\geq j\geq M-f$ the first term is:

$$\left[\underbrace{0\cdots0}_{1\times(p-1-j-M)}\quad G_{\Delta x}\quad\underbrace{0\cdots0}_{1\times(j-M+f)}\right](b_{p-1},...,b_0,...,b_{-f+1},b_{-f})'=\Delta Q_{\Delta x,j}b$$

Given the relation in (a.2):

$$\Delta Q_{\Delta x,j}b=\Delta Q_{\Delta x,j}D\widehat{B}=Q_{\Delta x,j}\widehat{B}$$

As long as $p-M\geq j\geq M-f$ the second term is the following:

$$\sum_{s=1}^{n}\left[\underbrace{0\cdots0}_{1\times(p-j-M)}\ G_{\Delta x,z_s}\ \underbrace{0\cdots0}_{1\times(j-M+f)}\right](\widehat{R}_{s,p},\widehat{R}_{s,p-1},...,\widehat{R}_{s,0},...,\widehat{R}_{s,-f+1},\widehat{R}_{s,-f})'=$$

$$\sum_{s=1}^{n}Q_{\Delta x,z_s,j}\widehat{R}_s$$

where

$$G_{\Delta x}=[\gamma_{\Delta x}(M),\ \gamma_{\Delta x}(M-1),\cdots,\gamma_{\Delta x}(0),\cdots,\gamma_{\Delta x}(M-1),\ \gamma_{\Delta x}(M)]$$

and

$$G_{\Delta x,z_s}=[\gamma_{\Delta x,z_s}(-M),\ \gamma_{\Delta x,z_s}(-M+1),\cdots,\gamma_{\Delta x,z_s}(0),\cdots,\gamma_{\Delta x,z_s}(M-1),\ \gamma_{\Delta x,z_s}(M)]\,,$$

$$s=1,...,n$$

Define similarly G_{z_r,z_s} and $G_{z_r,\Delta x}$ $r,s=1,...,n$. If $j>p-1-M$ the first term in (a.7) is:

$$\Delta Q_{\Delta x,j}b=\Delta Q_{\Delta x,j}D\widehat{B}=Q_{\Delta x,j}\widehat{B}$$

with

$$\Delta Q_{\Delta x,j} = \left[\overline{G}_{\Delta x}^{j-(p-1-M)} \quad \underbrace{0 \cdots 0}_{1\times(j-M+f)} \right]$$

where

$$\overline{G}_{\Delta x}^{j-(p-M-1)} = [\gamma_{\Delta x}(p-j-1)),\ \gamma_{\Delta x}(p-j), \cdots, \gamma_{\Delta x}(0), \cdots, \gamma_{\Delta x}(M-1),\ \gamma_{\Delta x}(M)]$$

that is, it is equal to $G_{\Delta x}$ with the first $j-(p-1-M)$ elements deleted. Also, as long as $j > p-M$ the second term in (a.7) is:

$$\sum_{s=1}^{n} Q_{\Delta x,z_s,j} \widehat{R}_s$$

with:

$$Q_{\Delta x,z_s,j} = \left[\overline{G}_{\Delta x,z_s}^{j-(p-M)} \quad \underbrace{0 \cdots 0}_{1\times(j-M+f)} \right]$$

and

$$\overline{G}_{\Delta x,z_s}^{j-(p-M)} = [\gamma_{\Delta x,z_s}(p-j)),\ \gamma_{\Delta x,z_s}(p-j+1), \cdots, \gamma_{\Delta x,z_s}(0), \cdots, \gamma_{\Delta x,z_s}(M-1),\ \gamma_{\Delta x,z_s}(M)]$$

Define similarly $Q_{z_r,z_s,j}$ and $Q_{z_r,\Delta x,j}$. Finally, if $j < M-f$ (a.7) is again:

$$Q_{\Delta x,j}\widehat{B} + \sum_{s=1}^{n} Q_{\Delta x,z_s,j} \widehat{R}_s$$

with

$$Q_{\Delta x,j} = \left[\underbrace{0 \cdots 0}_{1\times(p-1-j-M)} \quad \underline{G}_{\Delta x}^{M-f-j} \right] D$$

and

$$\underline{G}_{\Delta x}^{M-f-j} = [\gamma_{\Delta x}(M),\ \gamma_{\Delta x}(M-1), \cdots, \gamma_{\Delta x}(0), \cdots, \gamma_{\Delta x}(f+j-1),\ \gamma_{\Delta x}(f+j)]$$

that is, it is equal to $G_{\Delta x}$ with the last $M-f-j$ elements deleted. Also,

$$Q_{\Delta x,z_s,j} = \left[\underbrace{0 \cdots 0}_{1\times(p-j-M)} \; \underline{G}^{M-f-j}_{\Delta x,\Delta z_s} \right]$$

where

$$\underline{G}^{M-f-j}_{\Delta x,z_s} = [\gamma_{\Delta x,z_s}(-M),\ \gamma_{\Delta x,z_s}(-M+1),\cdots,\gamma_{\Delta x,z_s}(0),\cdots,\gamma_{\Delta x,z_s}(f+j-1),\ \gamma_{\Delta x,z_s}(f+j)]$$

Again, define similarly $Q_{z_r,\Delta z_s,j}$ and $Q_{z_r,\Delta x,j}$. Gather all the terms to get for $j=-f,...,p-1$

$$\begin{bmatrix} Q_{\Delta x,j} & Q_{\Delta x,z_1,j} & \cdots & Q_{\Delta x,z_n,j} \\ Q_{z_1,\Delta x,j} & Q_{z_1,j} & \cdots & Q_{z_1,z_n,j} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{z_n,\Delta x,j} & Q_{z_n,z_1,j} & \cdots & Q_{z_n,j} \end{bmatrix} \begin{bmatrix} \widehat{B} \\ \widehat{R}_1 \\ \vdots \\ \widehat{R}_n \end{bmatrix} = Q_j\widehat{W} \tag{a.8}$$

There are now only n equations missing. These are obtained in the first order conditions w.r.t. $W_p = (\widehat{R}_{1,p},...,\widehat{R}_{n,p})'$. Repeating arguments above these first order conditions are the following:

$$\int_{-\pi}^{\pi} [e^{-i\omega p} \begin{bmatrix} S_{z_1,\Delta x}(\omega) \\ \vdots \\ S_{z_n,\Delta x}(\omega) \end{bmatrix} \overline{B}(e^{i\omega})]d\omega =$$

$$\int_{-\pi}^{\pi} [e^{-i\omega p}\widetilde{S}_{\Delta x,z_1,...,z_n}(\omega) \begin{bmatrix} b(e^{i\omega}) \\ \widehat{R}_1(e^{i\omega}) \\ \vdots \\ \widehat{R}_n(e^{i\omega}) \end{bmatrix}]d\omega,$$

where $\widetilde{S}_{\Delta x,z_1,...,z_n}(\omega)$ is just the matrix $S_{\Delta x,z_1,...,z_n}(\omega)$ with the first row deleted. Using the results above we obtain:

$$\tilde{S}_p = \tilde{Q}_p\widehat{W}$$

where $\tilde{S}_p$ is just S_p as defined in (a.5) with the first element deleted and $\tilde{Q}_p$ is defined in the same way as the Q_j in (a.8) but with the first row deleted. We have finally $(p+f+1)\times(n+1)$ equations and unknowns:

$$V = Q\widehat{W} \tag{a.9}$$

where

$$V = \begin{bmatrix} S_{-f} & \cdots & S_{p-1} & \tilde{S}_p & 0 \end{bmatrix}'$$

and

$$Q = \begin{bmatrix} Q_{-f} \\ \vdots \\ Q_{p-1} \\ \tilde{Q}_p \\ U \end{bmatrix},$$

where

$$U = \begin{bmatrix} \underbrace{1 \;\; 1 \cdots 1}_{p+f+1} & \underbrace{0 \;\; 0 \cdots 0 \;\; 0 \;\; 0 \cdots 0}_{(p+f+1)\times n} \end{bmatrix}$$

The case when all the points in the series are used amounts to having p and f varying with t. Specifically, $p = t-1$ and $f = T-t$. Therefore, to get the weights that will be used to filter $\{x_t\}_{t=1}^{T}$ we need to solve the system in (a.9) T times. It is however easy to see that the L matrix in (a.9) will always be the same, it does not vary with t. Only the V vector varies with t. An algorithm in pseudo-code that constructs all the objects needed to solve the problem is presented in Appendix B.

A.1: The stationary case

The case when x_t is covariance stationary amounts to straightforward modifications of the formulae presented above. First, the restriction $\widehat{B}(1) = 0$ is not necessary. The derivation is very similar and results again in a linear system with $(p+f+1)\times(n+1)$ equations and unknowns:

$$V = Q\widehat{W}$$

where

$$V = \begin{bmatrix} S_{-f} & \cdots & S_{p-1} & S_p \end{bmatrix}'$$

with the S_j, $j = -f, ..., p$ again as defined in (a.5) and

$$Q = \begin{bmatrix} Q_{-f} \\ \vdots \\ Q_p \end{bmatrix},$$

where the Q_j, $j = -f, ..., p-1$ are as defined in (a.8), but with reference to the series $x, z_1, ..., z_n$ and not $\Delta x, z_1, ..., z_n$ as was the case before. Additionally, Q_p is also as defined in (a.8) for $j = p$.

Appendix B: Pseudo-Code Algorithm to compute filtered series

Case when all the observations are used, so that $p = t-1$ and $f = T-t$

Inputs:

Series - Series of interest (x_t) with linear trend removed

Indicators - list of demeaned indicators $\{z_1, ..., z_n\}$, where $z_i = \{z_{i,1}, z_{i,2}, ...z_{i,T}\}$,

$i = 1, ..., n$. T is the sample size

lowerperiod - lower period in the band of interest (must be ≥ 2)

upperperiod - upper period in the band of interest

M - truncation point in the estimation of the spectrum

Auxiliary functions used:

AppendTo[*list*, *element*] returns *list* with *element* appended

$Join[listA, listB]$ returns a list formed by appending $listB$ to $listA$

$Drop[list, s]$ returns $list$ with the first s elements dropped

$Drop[list, -s]$ returns $list$ with the last s elements dropped

Note: Matrices are represented as a list of lists. If A is a matrix the element i, j is accessed by $A[i][j]$.

Algorithm:

Step 1

Calculate:

$upperfrequency = 2\pi/lowerperiod$ (*upper frequency in the band*)

$lowerfrequency = 2\pi/upperperiod$ (*lower frequency in the band*)

$autocovariance[k] = \frac{1}{T}(1 - \frac{k}{M+1})\sum_{t=k+1}^{T}((\Delta x_t, z_{1,t}, ..., z_{n,t})')((\Delta x_{t-k}, z_{1,t-k}, ..., z_{n,t-k})')'$

(*estimated autocovariance at lag k with Bartlett Kernel, $k = 1, ..., M$*)

Create $B = \{(upperfreq - lowerfreq)/\pi, (\sin(upperfreq) - \sin(lowerfreq))/\pi,$

$(\sin(2.upperfreq) - \sin(2.lowerfreq))/2\pi, ...,$

$(\sin((T-1+M)upperfreq) - \sin((T-1+M)lowerfreq))/(T-1+M)\pi\}$

(*list with ideal filter weights*)

For $j = 1 - T$ to $j = T - 1$:

Create $V_j = \{B[|j|+1]autocovariance[0][1][1]$
$+\sum_{i=1}^{M}(B[|j|+i+1] + B[||j|-i|+1])autocovariance[i][1][1]\}$

  For $s = 2$ to $s = n + 1$:

  $AppendTo[V_j, B[|j|+1]autocovariance[0][s][1]$
  $+\sum_{i=1}^{M}(B[|j-i|+1]autocovariance[i][s][1] + B[j+i|+1]autocovariance[-i][1][s])]$;

  End cycle

End cycle

Step 2

For $j = T - 1 - M$ to $j = T - 1 + M$ get numerically:

$$lhsIntegr[j] = \int_{-upperfrequency}^{-lowerfrequency} [e^{i\omega j}\frac{1}{1-e^{i\omega}}]d\omega +$$

$$\int_{lowerfrequency}^{upperfrequency} [e^{i\omega j}\frac{1}{1-e^{i\omega}}]d\omega$$

End of cycle

Create $S_{1-T} = \{\}$;

For $j = 1$ to $j = n+1$

$AppendTo[S_{1-T}, \frac{1}{2\pi}(autocovariance[0][1][j] \times lhsIntegr[T-1]$
$+\sum_{i=1}^{M}(autocovariance[i][1][j] \times lhsIntegr[T-1-i]) + autocovariance[i]'[1][j] \times$
$lhsIntegr[T-1+i])]$

End cycle

For $j = 1 - T$ +1 to $j = T-1$ calculate $S_j = S_{j-1} + V_j$

End cycle

Step3 (* Build matrix Q*)

Create $D = \{\}$;

For $j = 1$ to $j = T-1$ (*Create matrix D*)

$AppendTo[\ D\ ,Join[\{-1,-1,...,-1\}_{(1\times j)},\{0,0,...,0\}_{(1\times(T-j))}]]$

End cycle

Create $zerosList = \{0,0,...,0\}_{(1\times(T-M-2))}$

For $i = 1$ to $i = n+1$

For $j = 1$ to $j = n+1$

$BlockRow[i][j] = \{autocovariance[-M][i][j], ..., autocovariance[0][i][j], ...,$
$autocovariance[M][i][j]\}$

$BlockRow[i][j] = Join[AppendTo[zerosList, 0], BlockRow[i][j]]$;

$BlockRow[i][j] = Join[BlockRow[i][j], zerosList]$;

End cycle

End cycle

Create $Q = \{\}$;

set $r = 0$;

Create $Q[r] = \{\}$;

For $i = 2$ to $i = n + 1$,

$rowk = Append[Drop[BlockRow[k][1], T - 1 - r + 1], 0].D$;

For $j = 2$ to $j = n + 1$,

$rowk = Join[rowk, Append[Drop[BlockRow[k][p], T - 1 - r], 0]]$;

$AppendTo[Q[r], rowk]$;

End cycle

End cycle

$Q = Join[Q, Q[r]]$;

For $r = 1$ to $r = T - 1$

$Q[r] = \{\}$;

For $k = 1$ to $k = n + 1$

$rowk = Drop[Drop[BlockRow[k][1], T - 1 - r + 1], -(r - 1)].D$;

For $p = 2$ to $p = n + 1$

$rowk = Join[rowk, Drop[Drop[BlockRow[k][p], T - 1 - r], -(r - 1)]]$

End cycle

$AppendTo[Q[r], rowk]$;

End cycle

$Q = Join[Q, Q[r]]$

End cycle

$lastMat = \{\}$;

$AppendTo[lastMat, Join[\{1, 1, ..., 1\}_{(1\times T)}, \{0, 0, ..., 0\}_{(1\times(T\times n))}]]$;

$Q = Join[Q, lastMat]$;

Step4 (*Create vector V for each observation, compute weights and filter the data*)

$zeroslist = \{0\};$

$filteredData = \{\};$

For $t = 1$ to $t = T$

$p = t - 1;$

$f = T - t;$

$V = \{\};$

For $c = p - 1$ to $c = -f$ (c decreasing)

$V = Join[V, S_c]$

End cycle

$V = Join[Drop[S_p, 1], V];$

$V = Join[V, zeroslist];$

Solve the linear system $Qsolution = V$ w.r.t. $solution$

$AppendTo[filtereddata, \sum_{j=-f}^{p} solution[T - j - f] \times Series[t - j] +$
$\sum_{k=1}^{n} \sum_{j=-f}^{p} solution[(T - j - f) + T \times k] \times Indicators[k][t - j]]$

End cycle

Return $filtereddata$

Appendix C: Data

The sample runs from the first quarter of 1967 to the second quarter of 2005. In addition to real U.S. GDP, the following indicators were selected:

Industrial Production Index (IPI), monthly series - quarterly series is constructed as average of the three months of each quarter. Available from the Board of Governors of the Federal Reserve system. Series ID: INDPRO

Capacity Utilization (Total Industry), monthly series - quarterly series is constructed as average of the three months of each quarter. Available from the Board of Governors of the

Federal Reserve system. Series ID: TCU

Non-Farm Output, quarterly series, seasonally adjusted- Available from the U.S. Department of Labor: Bureau of Labor Statistics. Series ID:OUTNFB

Business Sector: Hours of All Persons, quarterly series, seasonally adjusted. Available from the U.S. Department of Labor: Bureau of Labor Statistics. Series ID:HOABS

Average weekly hours, monthly series, seasonally adjusted. quarterly series is constructed as average of the three months of each quarter. Available from the U.S. Department of Labor: Bureau of Labor Statistics. Series ID:AWHNONAG

Help Wanted Adds Index, monthly index, seasonally adjusted - quarterly series is constructed as average of the three months of each quarter. Available from the Conference Board. Series ID:HELPWANT

All data are in logarithms. First differences are applied to all the indicators.

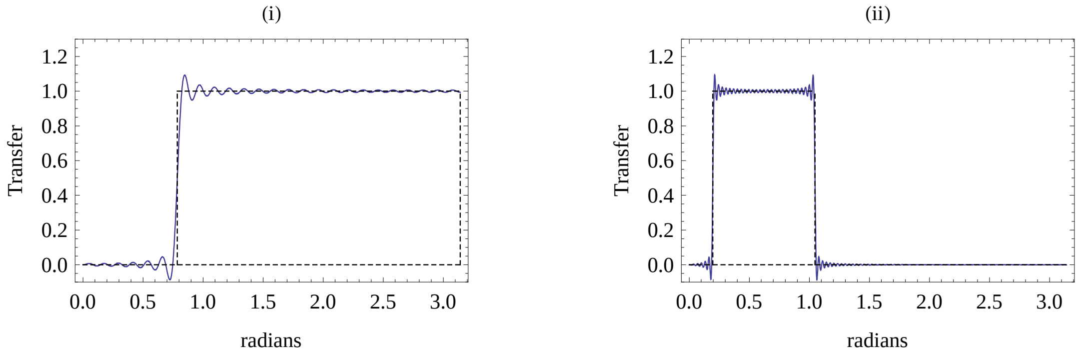

Figure 1: Transfer function of the BK filter: (i) $p = f = 50$ with [2,8] periods band (ii) $p = f = 200$ with [6,32] periods band. The dashed lines represent the ideal filters that isolate the referred bands.

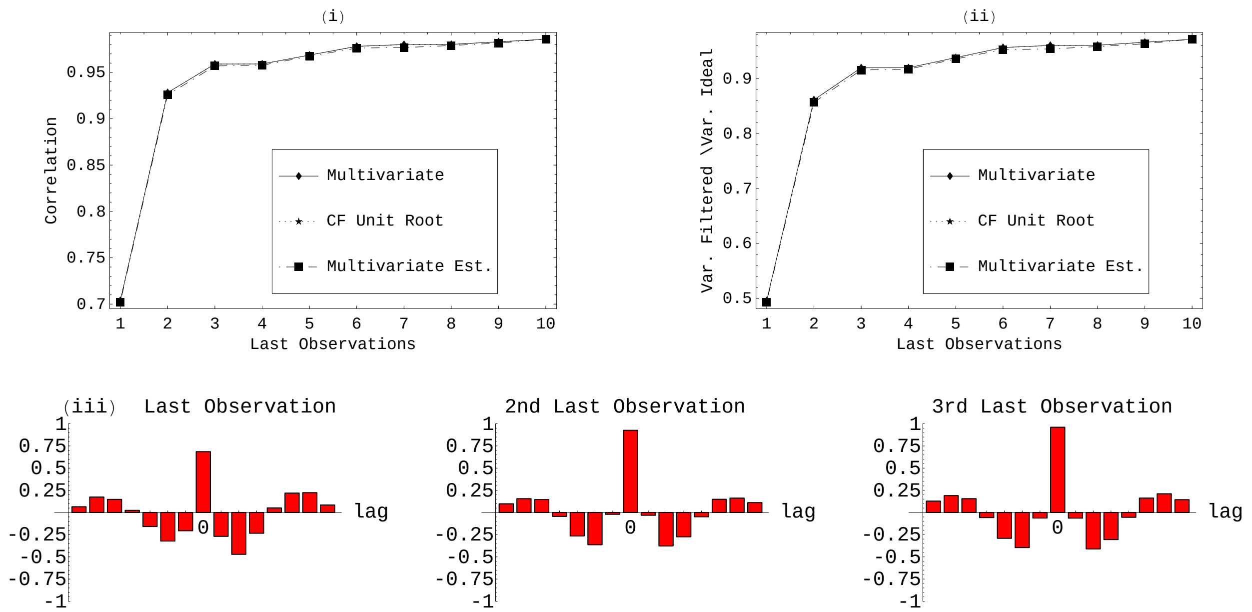

Figure 2: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Coincident-Low Corr.-3 Indicators*, [2, 8] periods band. Here and in the remainder of the analysis, 1 in the horizontal axis of (i) and (ii) represents the last observation, 2 refers to the second last observation and so forth.

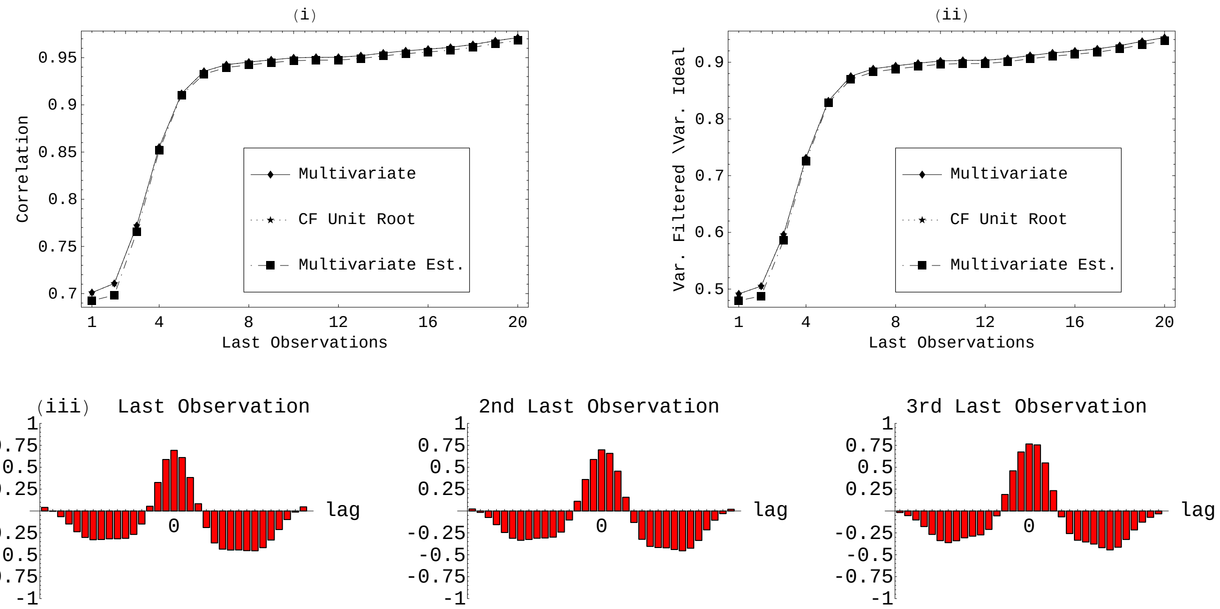

Figure 3: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Coincident-Low Corr.-3 Indicators*, [6, 32] periods band.

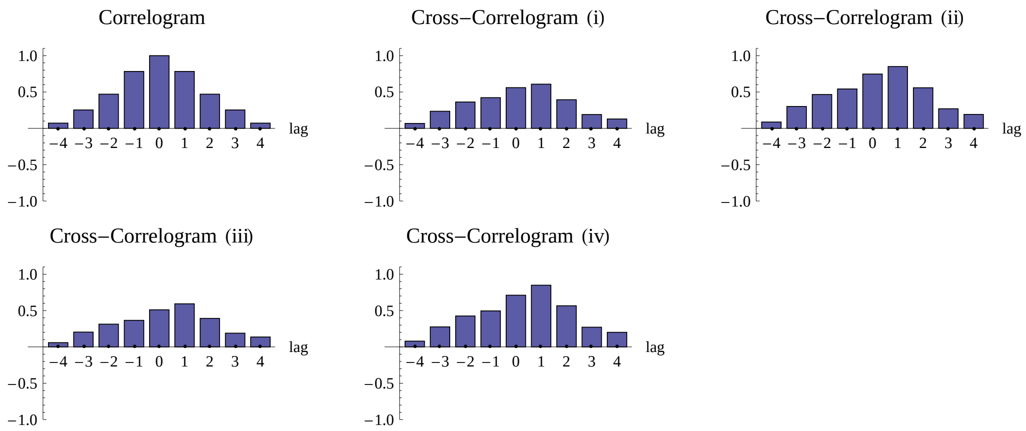

Figure 4: Second moments for DGP - Correlogram of Δx_t and cross-correlogram of $(\Delta x_t, z_{i,t})$: (i) *Leading-Low Corr.-3 Indicators*,(ii) *Leading-High Corr.-3 Indicators*, (iii) *Leading-Low Corr.-10 Indicators*, (iv) *Leading-High Corr.-10 Indicators*. Annual Data.

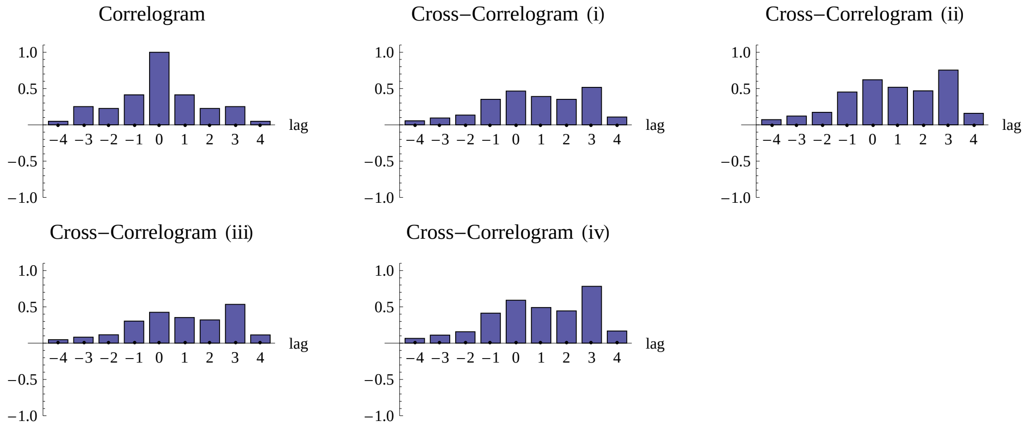

Figure 5: Second moments for DGP - Correlogram of Δx_t and cross-correlogram of $(\Delta x_t, z_{i,t})$:(i) *Leading-Low Corr.-3 Indicators*,(ii) *Leading-High Corr.-3 Indicators*, (iii) *Leading-Low Corr.-10 Indicators*, (iv) *Leading-High Corr.-10 Indicators*. Quarterly Data.

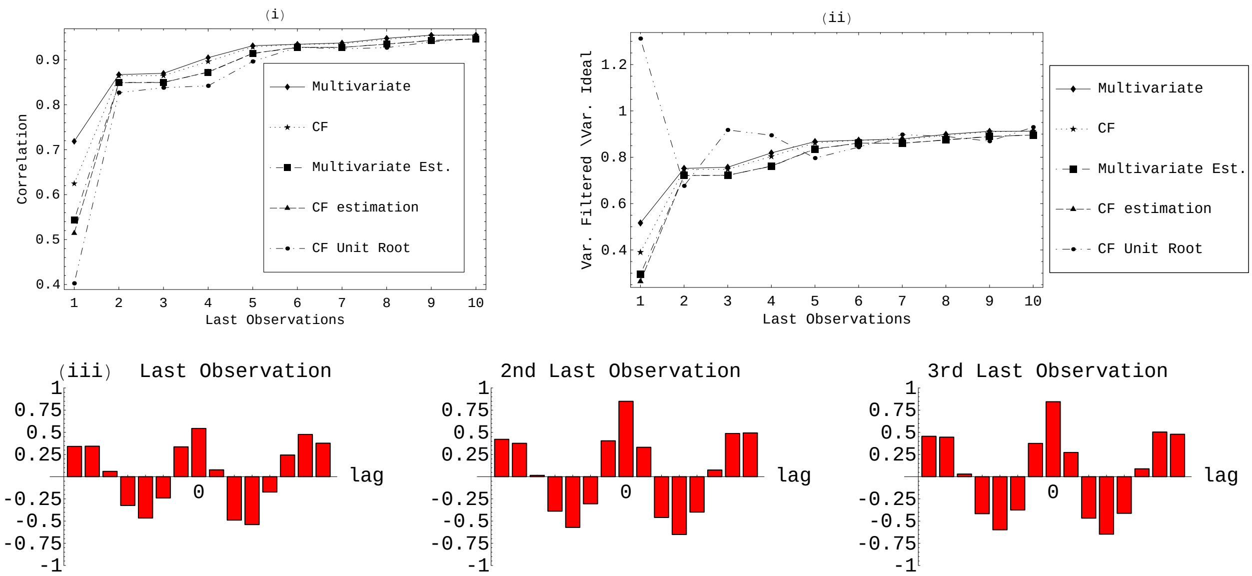

Figure 6: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Leading-Low Corr.-3 Indicators*, [2, 8] periods band.

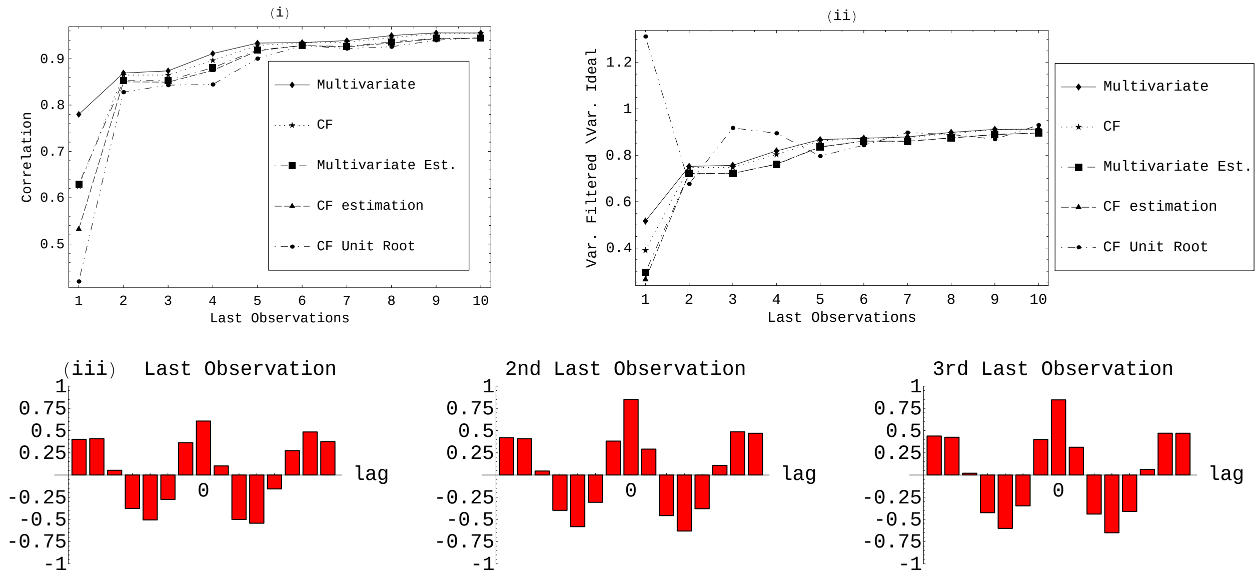

Figure 7: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Leading-High Corr.-10 Indicators*, [2, 8] periods band.

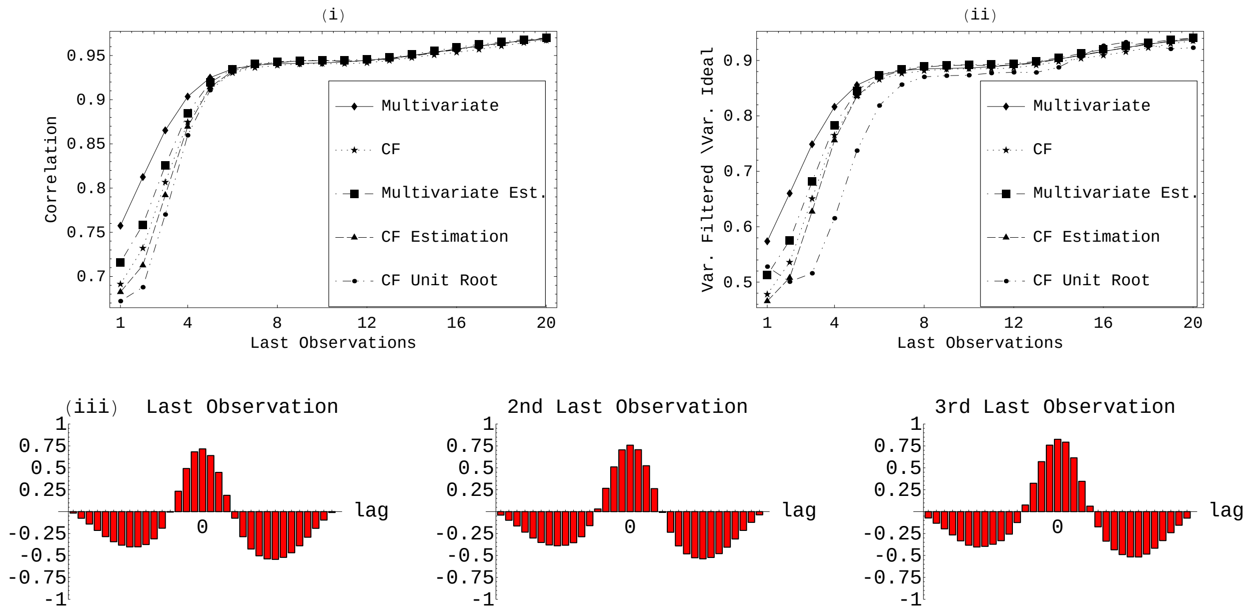

Figure 8: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Leading-Low Corr.-3 Indicators*, [6, 32] periods band.

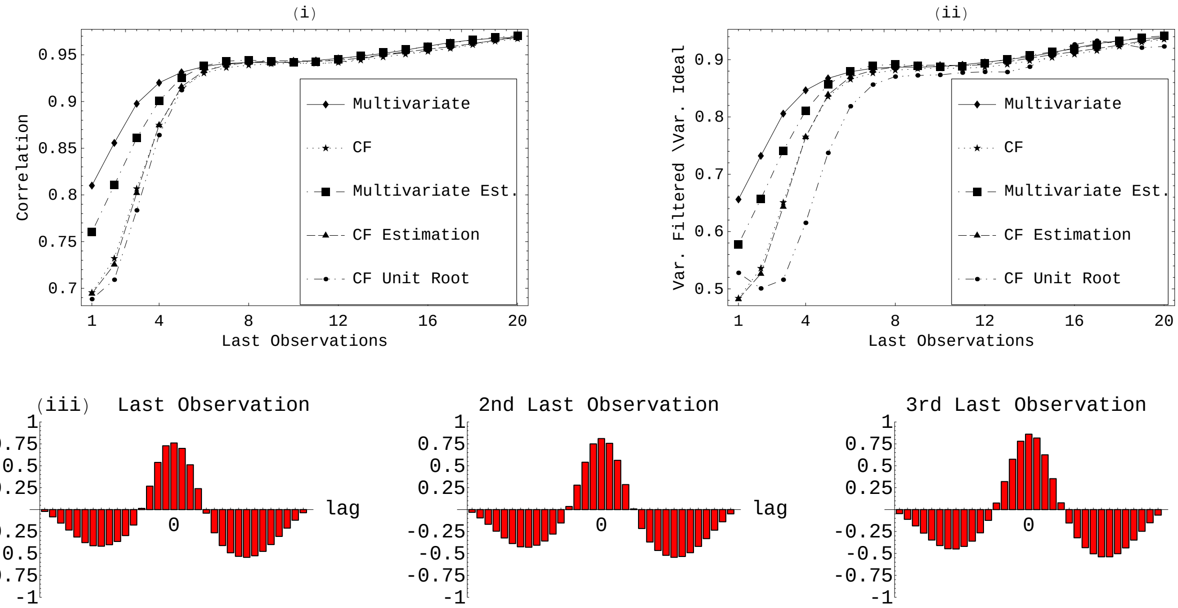

Figure 9: (i) Correlation with ideally filtered series, (ii) variance as proportion of the variance of the ideally filtered series and (iii) cross-correlation function between *Multivariate* filtered and ideally filtered series. *Leading-High Corr.-10 Indicators*, [6, 32] periods band.

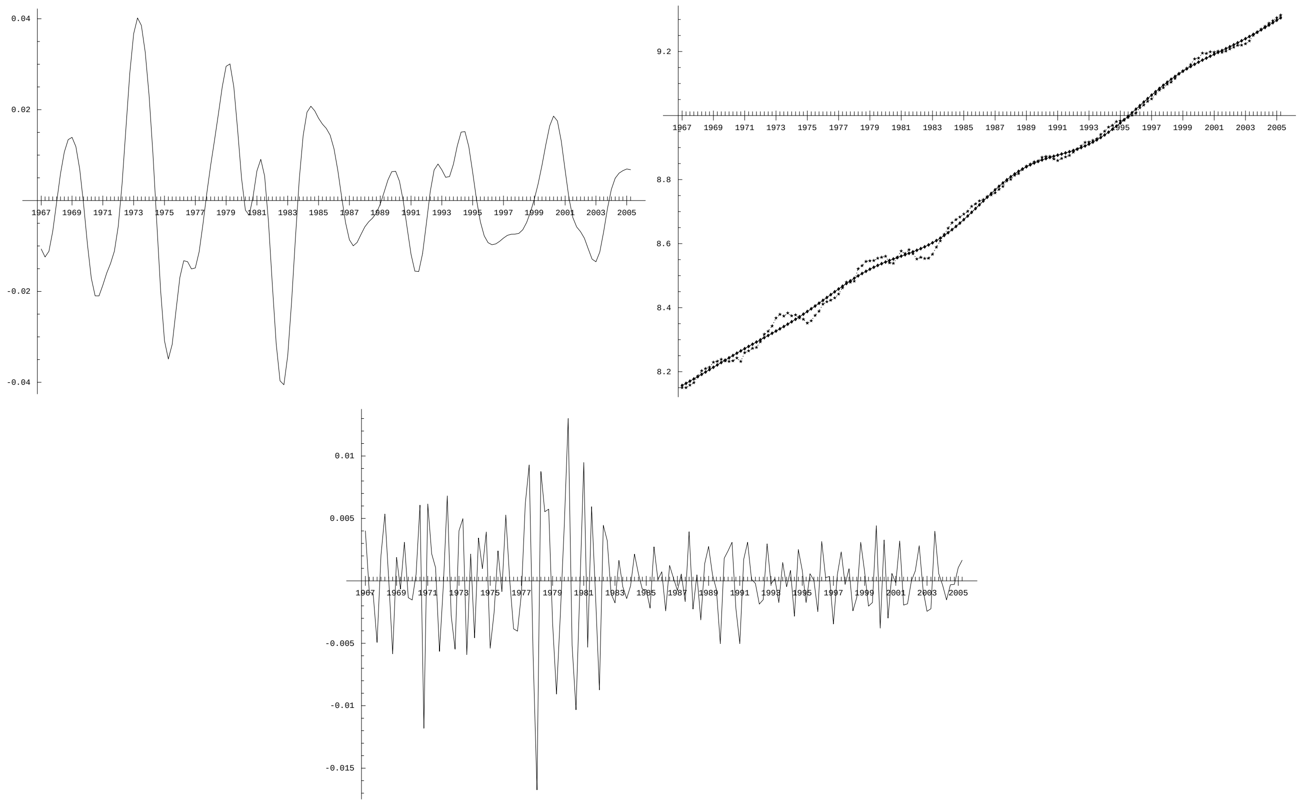

Figure 10: Decomposition of U.S. quarterly GDP into cycle, trend and noise components (1967-1 to 2005-2).

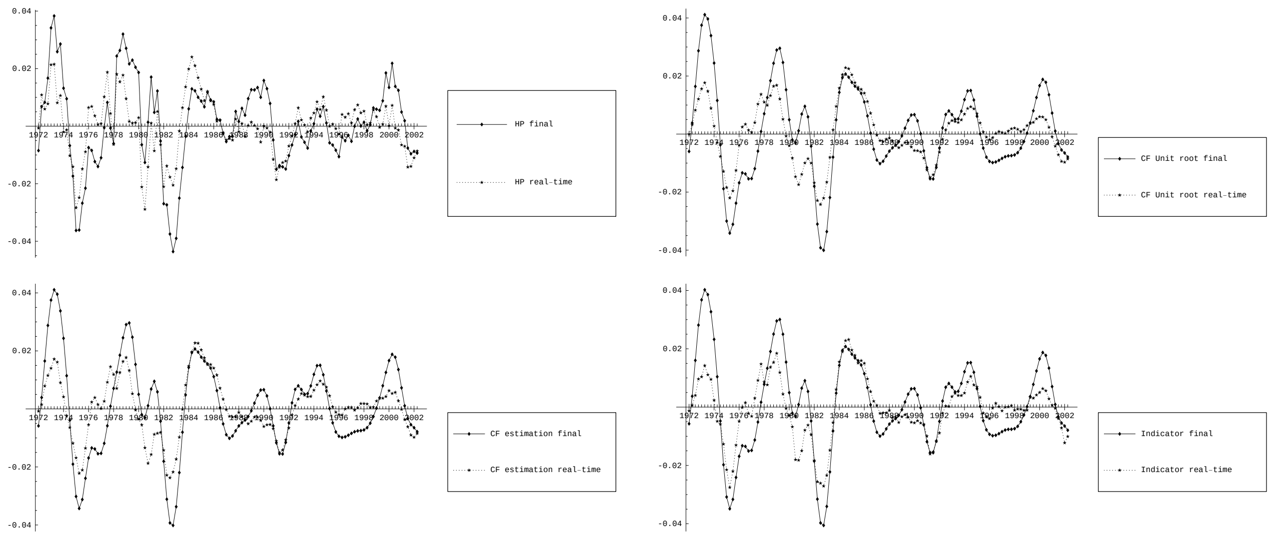

Figure 11: Real-time and final estimates of business cycle component using the HP filter, the CF filter assuming that GDP is a random walk, the CF filter with estimated second moments (truncation point M=6) and the Multivariate filter with estimated second moments (M=6)(1972-1 to 2002-2).

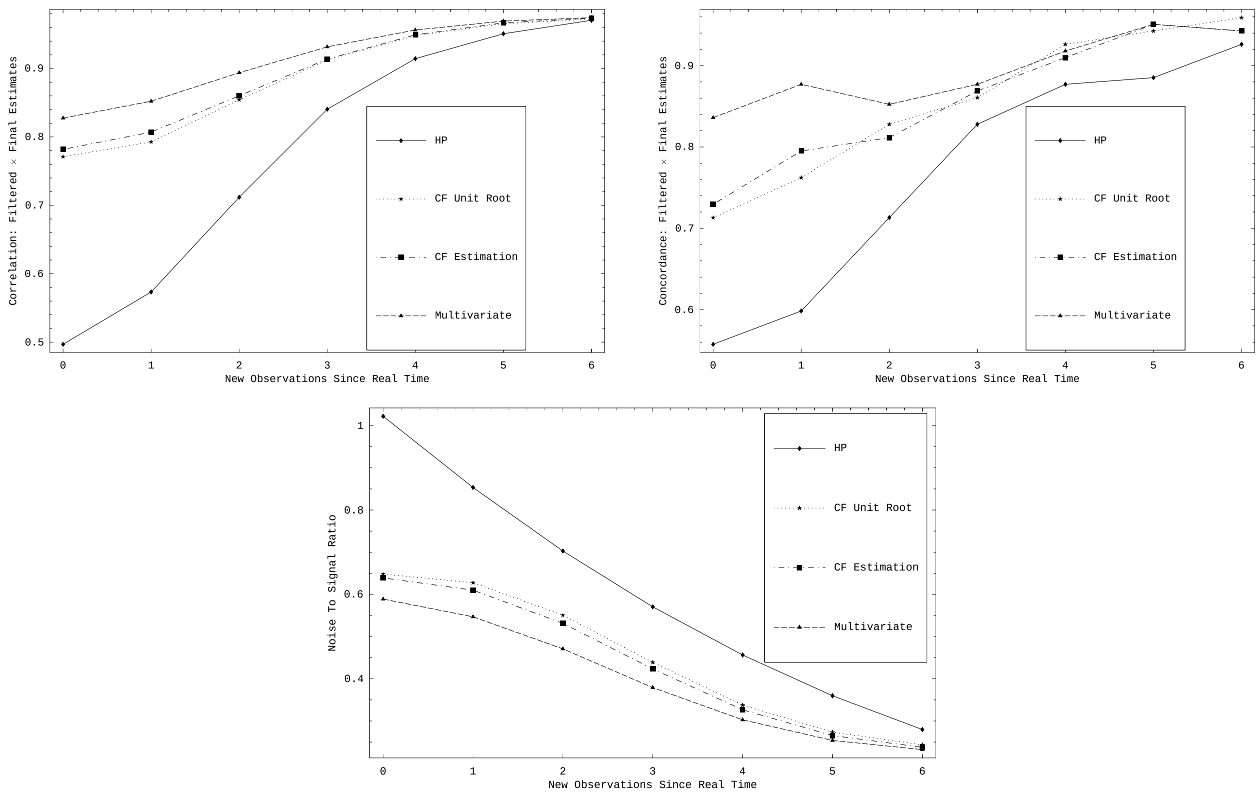

Figure 12: Correlation, Sign Concordance and Noise To Signal Ratio: Final versus Real-time estimates when new observations are available (1972-1 to 2002-2).

References

[1] Baxter, M. and King, R. (1999). "Measuring business cycles: approximate band-pass filters for economic time series". *Review of Economics and Statistics*, 81:575-93.

[2] Brockwell, P.J. and Davis, R. (1991). *Time Series: Theory and methods* (2nd ed.). Springer

[3] Bujosa A., Bujosa, M. and García-Ferrer (2001). "A Note on Pseudo-spectra and Pseudo-Covariance Generating Functions of ARMA processes". mimeo

[4] Christiano, L. and T. Fitzgerald (2003)."The band-pass filter". *International Economic Review*, 44:435-65.

[5] Geweke, J. (1978). "The revision of seasonally adjusted time series". In Proceedings of the Business and Economic Statistics Section, American Statistical Association, 320–325.

[6] Gomez, V. (2001). "The use of Butterworth filters for trend and cycle estimation in economic time series". *Journal of Business and Economic Statistics*, 19:365-73.

[7] Harvey, A.C. (1993). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press

[8] Harvey, A. C. and T. Trimbur (2003). "Generalised model-based filters for extracting trends and cycles in economic time series". *Review of Economics and Statistics*, 85:244-55.

[9] Hodrick, R. J. and E. C. Prescott (1997). "Postwar U.S. business cycles: an empirical investigation". *Journal of Money, Credit and Banking*, 29:1-16.

[10] Hurvich, C.M., Ray, B.K., (1995). "Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes". *Journal of Time Series Analysis*, 16:17-42.

[11] Kaiser, R. and Maravall, A. (1999). "Estimation of the Business cycle: A Modified Hodrick-Prescott Filter", Banco de España -Servicio de Estudios, Working Paper 9912

[12] King, R. and Rebelo, S. (1993). "Low Frequency Filtering and Real Business Cycles", *Journal of Economic Dynamics and Control*, 17:207-231.

[13] Orphanides, A. and S. van Norden (2002). "The unreliability of output-gap estimates in real time". *Review of Economics and Statistics*, 84:569-83.

[14] Pedersen, T. M.(2001). "The Hodrick–Prescott filter, the Slutzky effect, and the distortionary effect of filters". *Journal of Economic Dynamics and Control*, 25:1081-1101.

[15] Pierce, DA. (1980). "Data revisions with moving average seasonal adjustment procedures". *Journal of Econometrics*, 14:95-114.

[16] Priestley, M.B. (1981). *Spectral Analysis of Time Series*, Academic Press, London

[17] Rünstler, G. (2004). "Modelling phase shifts among stochastic cycles". *Econometrics Journal*, 7:232-248.

[18] Stock, J. and M. W. Watson (1999). Business cycle fluctuations in US macroeconomic time series. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, pp. 3-64. Amsterdam: Elsevier Science Publishers.

[19] Valle e Azevedo, J., Koopman, S.J. and Rua, A (2006). "Tracking the business cycle of the Euro area: A Multivariate model based band-pass filter". *Journal of Business and Economic Statistics*, 24(3):278-290

[20] Velasco, C., (1999). "Non-stationary log-periodogram regression". *Journal of Econometrics*, 91:325-371.