TRACKING THE US BUSINESS CYCLE WITH A SINGULAR SPECTRUM ANALYSIS

Miguel de Carvalho† , Paulo C. Rodrigues†, António Rua†

Abstract

The monitoring of economic developments is an exercise of considerable importance for policymakers, namely, central banks and fiscal authorities as well as for other economic agents such as financial intermediaries, firms and households. However, the assessment of the business cycle is not an easy endeavor as the cyclical component is not an observable variable. In this paper we resort to singular spectrum analysis in order to disentangle the US GDP into several underlying components of interest. The business cycle indicator yielded through this method is shown to bear a resemblance with band-pass filtered output. As the end-of-sample behavior is typically a thorny issue in business cycle assessment, a real-time estimation exercise is here conducted to assess the reliability of the several filters. The obtained results suggest that the business cycle indicator proposed herein possesses a better revision performance than other filters commonly applied in the literature.

KEYWORDS: Band-Pass Filter; Principal Components; Singular Spectrum Analysis; US Business Cycle.

JEL CLASSIFICATION: C50, E32.

† Miguel de Carvalho (Email: *mb.carvalho@fct.unl.pt*) and Paulo C. Rodrigues (Email: *paulocanas@fct.unl.pt*) are full researchers at CMA – *Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa*, Portugal. António Rua (Email: *antonio.rua@bportugal.pt*) is head of the Conjuntural Unit of the Forecasting and Conjuntural Division of the Economics Research Department of *Banco de Portugal*. This paper was written during a visit of Miguel de Carvalho to *Banco de Portugal*. The usual disclaimer applies.

1. INTRODUCTION

Since the seminal work of Burns and Mitchell (1946), the question of how to extract the latent cyclical component of an economic series has been a subject of considerable attention. The wealth of research resources deposited in the analysis of business cycles is certainly a corollary of its importance for a broad audience of decision makers including central banks, fiscal authorities, financial intermediaries, firms and households. Among the most routinely employed methods to extract the cyclical component of an economic series lies the filter proposed by Hodrick and Prescott (1997).[1] Despite of its popularity and its wide range of applicability, the filter proposed by Hodrick and Prescott is a high-pass filter thus incorporating excessive irregular noisy movements in the cyclical component. In opposition to this high-pass filter, Baxter and King (1999) advocate the use of a band-pass filter, i.e., a filter which suppresses a low frequency component and the high frequency fluctuations in a time series. As advocated by the latter authors, the seminal definition of business cycle advanced by Burns and Mitchell is most congruous with the technology of band-pass filtering than with high-pass methods, focusing on components whose recurring movements range from 6 to 32 quarters. Regardless of its appealing features the aforementioned band-pass filter lacks the ability to produce end-of-sample estimates, which are particularly precious for policymaking purposes. An optimal finite sample approximation to the ideal band-pass filter, which is able to cope with the problem of end-of-sample estimates was proposed by Christiano and Fitzgerald (2003). As Christiano and Fitzgerald put forward, the potential misspecification of the true data generating process as a random walk can still render nearly optimal performance. Even though the three methods mentioned above are certainly the

[1] For an inventory of applications of this filter see, for instance, Ravn and Uhlig (2002), and references therein.

most employed in applications, the diversity of the literature includes many other possibilities.[2] For example, Yogo (2008) makes use of multiresolution wavelet analysis in order to decompose the US GDP into several components of interest. The wavelet-filtered data is then shown to behave likewise the band-pass filtered series. Pollock (2000) mentions the lack of ability of some of the most widely used filters for coping with commonly observed attributes of economic time series (say, a structural break) and suggests the use of the rational square wave filter – a method known to engineers as the digital Butterworth filter. As a means to take full advantage of the information available in other variables related with economic activity, Azevedo, Koopman and Rua (2006) proposed a multivariate band-pass filter using a state-space approach.

In this paper we employ singular spectrum analysis in order to disentangle the US GDP into several underlying components of interest. Singular spectrum analysis is a natural extension of principal components analysis for time series data, with known applications in climatology (Allen and Smith, 1996), geophysics (Kondrashov and Ghil, 2006), as well as meteorology (Paegle, Byerle and Mo, 2000). Other recent applications of singular spectrum analysis include, for example, forecasting (Hassani, Heravi and Zhigljavsky, 2009). For a comprehensive overview of singular spectrum analysis see, for instance, Golyandina, Nekrutkin and Zhigljavsky (2001). The business cycle indicator yielded by dint of this method is shown to bear a resemblance with band-pass filtered output and is broadly in line with the contraction and expansion periods dated by the US Business Cycle Dating Committee of the National Bureau of Economic Research (NBER).

From a policymaking standpoint a feature of remarkable importance is the reliability of real-time estimates of the business cycle indicator (see, for instance, Orphanides and van

[2] At the time of writing, according to Google Scholar, the papers of Hodrick and Prescott (1997), Baxter and King (1999) and Christiano and Fitzgerald (2003) hold 3014, 1541 and 487 citations, respectively.

Norden, 2002). However, the end-of-sample behavior of business cycle indicators is a thorny issue shared by several frequently applied filters. Indeed, one should bear in mind that it is extremely difficult to estimate output gap in real time, given that only with time one can be more precise about the "true" position at a given period. In order to compare the revisions of the business cycle indicator proposed in this paper and the remainder filters, a real-time exercise is here conducted. Our results indicate that the business cycle indicator proposed herein possesses a better revision performance than other filters commonly employed in the literature.

The structure of this paper is as follows. The next section introduces singular spectrum analysis. Section 3 is devoted to the extraction of the cyclical component of the US GDP. Summary and conclusions are given in Section 4.

2. SINGULAR SPECTRUM ANALYSIS IN A NUTSHELL

2.1 Prefatory Decomposition Theory

To lay the groundwork, below we set forth some considerations on the orthogonal representation of stochastic processes. These are related with the theoretical underpinning which underlies the basic motivation behind singular spectrum analysis.

One of the most widely known orthogonal representation of a stochastic process is given by the Karhunen-Loève decomposition (Loève, 1978), which is stated in Lemma 1 below. Roughly speaking this decomposition guarantees that any random variable which is continuous in quadratic mean can be represented as a linear combination of orthogonal functions. The proof of this result can be found elsewhere (*e.g.* Loève (1978), pp. 144–145).

Lemma 1 *A random function $Y(t)$ which is continuous in quadratic mean on a closed interval $I = [0, t]$ admits on I an orthogonal decomposition of the form*

$$Y(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\omega_i(t)\nu_i, \tag{1}$$

for some stochastic orthogonal quantities ν_i, iff λ_i and $\omega_i(t)$ respectively denote the eigenvalues and the orthonormalized eigenfunctions of the autocovariance function $\Upsilon(r, s)$.

Despite the broadness of this theoretical result, in practice a discrete variant of decomposition (1) is often preferred to conduct data analysis. Hence, in lieu of considering the eigenfunctions one uses instead the eigenvectors of a discrete version of the autocovariance function $\Upsilon(r, s)$. Additionally, from the practical stance, one is often confined to the truncation of a finite number of terms in the decomposition (1). Indeed, the name of this decomposition is at the origin of the alternative naming of principal component analysis as Karhunen-Loève transformation.[3]

2.2 Modus Operandi of Singular Spectrum Analysis

In this section we briefly describe the *modus operandi* of singular spectrum analysis. A primer on the methods described below can be found for instance in Golyandina, Nekrutkin and Zhigljavsky (2001). In essence, the crux of the method can be dissociated in two phases, namely decomposition and reconstruction. Each of these phases includes two steps. The phase of decomposition includes the steps of embedding and singular value decomposition, which we introduce below.

Embedding

This is the preparatory step of the method. The core concept assigned to this step is given by the trajectory matrix, i.e., a lagged version of the original time series $\mathbf{y} = [y_1 \ \cdots \ y_n]'$. Formally, the trajectory matrix is defined as

[3] Further details between connections of principal component models and Karhunen-Loève decomposition can be found, for example, in Basilevsky and Hum (1979).

$$\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_\kappa \\ y_2 & y_3 & \cdots & y_{\kappa+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_l & y_{l+1} & \cdots & y_{l+(\kappa-1)} \end{bmatrix}, \tag{2}$$

where κ is such that $\mathbf{Y}$ encompasses all the observations in the original time series, i.e., $\kappa = n - l + 1$.

In order to set terminology, we refer to each vector $\mathbf{y}_i = [y_i \ \cdots \ y_{l+(i-1)}]'$, $i = 1, \ldots, \kappa$, as a window. The window length l, is a parameter to be defined by the user. Observe that $\mathbf{Y}$ is a Hankel matrix, where the original series $\mathbf{y}$ relies in the junction formed by the first column and the last row. It can also be useful to think of the trajectory matrix as a sequence of κ windows, i.e., $\mathbf{Y} = [\mathbf{y}_{1,l} \ \cdots \ \mathbf{y}_{\kappa,l}]$.

The trajectory matrix also finds application in the estimation of the lag-covariance matrix $\mathbf{\Sigma}$ of the source series $\mathbf{y}$. In effect, Broomhead and King (1986) proposed the following estimate

$$\widehat{\mathbf{\Sigma}} = \frac{1}{l}\mathbf{Y}'\mathbf{Y}. \tag{3}$$

As an alternative, one can also rely the estimate of the lag-covariance matrix using the proposal of Vautard and Ghil (1989), which is given by

$$\widehat{\mathbf{\Sigma}} = \left[\frac{1}{n-|i-j|}\sum_{t=1}^{n-|i-j|} y_t y_{t-|i-j|}\right]_{l\times l}. \tag{4}$$

Observe that this estimate yields a Toeplitz matrix, i.e., a diagonal-constant matrix.

Singular Value Decomposition

In the second step we perform a singular value decomposition of the trajectory matrix. Hence, from a eigenanalysis of the matrix $\mathbf{Y}\mathbf{Y}'$ we collect the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_d$, where $d = rank(\mathbf{YY}')$, as well as the corresponding left and right singular vectors which we respectively denote by $\mathbf{w}_i$ and $\mathbf{v}_i$. Thus, we are able to rewrite the trajectory matrix as

$$\mathbf{Y} = \sum_{i=1}^{d} \sqrt{\lambda_i} \mathbf{w}_i \mathbf{v}_i', \tag{5}$$

which somehow resembles equation (1).

We now turn to the second phase of the method – reconstruction. This includes the steps of grouping and diagonal averaging.

Grouping

In the grouping step, the selection of the m principal components takes place. Formally, let $I = \{1, \ldots, m\}$ and $I^c = \{m+1, \ldots, d\}$. The point here is to choose the first m leading eigentriples associated to the signal and exclude the remaining $(d - m)$ associated to the noise. Stated differently, at the core of this step we have the proper selection of the set I, as a means to disentangle the series $\mathbf{Y}$ into

$$\mathbf{Y} = \sum_{i \in I} \sqrt{\lambda_i} \mathbf{w}_i \mathbf{v}_i' + \boldsymbol{\varepsilon}, \tag{6}$$

where $\boldsymbol{\varepsilon}$ denotes an error term, and the remainder summands represents the signal. In practice, this is typically done through readjusted methods for selecting a reasonable number of principal components m.

Diagonal Averaging

The central idea in this step is the reconstruction of the deterministic component of the series. A natural way to do this is to transfigure the matrix $\mathbf{Y} - \boldsymbol{\varepsilon}$ obtained in the previous step, into an Hankel matrix. The point here is to reverse the process done so far, returning to a reconstructed variant of the trajectory matrix (2), and thus the deterministic component of the series. An optimal way to do this is to average over all the elements of the several antidiagonals. Formally, consider the linear space $\mathcal{M}_{l,\kappa}$ formed by the collection of all the $l \times \kappa$ matrices, and let $\{\mathbf{h}_l\}_{l=1}^{n}$ denote the canonical basis of $\mathbb{R}^n$. Further, consider the matrix $\mathbf{X} = [x_{i,j}] \in \mathcal{M}_{l\times\kappa}$. The diagonal averaging procedure is hence carried on by the mapping $\overline{\mathbb{D}} : \mathcal{M}_{l\times\kappa} \to \mathbb{R}^n$ defined as

$$\overline{\mathbb{D}}(\mathbf{X}) = \sum_{w=2}^{\kappa+l} \mathbf{h}_{w-1} \sum_{(i,j)\in\mathcal{A}_w} \frac{x_{i,j}}{|\mathcal{A}_w|}. \tag{7}$$

Here $|.|$ stands for the cardinal operator, and

$$\mathcal{A}_w = \{(i,j) : i + j = w\}, \text{ for } i = 1, \ldots, l, j = 1, \ldots, \kappa. \tag{8}$$

Hence we are now able to write the deterministic component of the series through the diagonal averaging procedure described above

$$\widetilde{\mathbf{y}} = \overline{\mathbb{D}}\left(\sum_{i\in I} \sqrt{\lambda_i}\mathbf{w}_i\mathbf{v}_i'\right). \tag{9}$$

Here, the tildes are used to denote reconstruction.

3. MEASURING THE US BUSINESS CYCLE

3.1 Tracking the Business Cycle

In the sequel we construct a business cycle indicator through the decomposition method introduced in the foregoing section.[4]

Data from the quarterly US GDP were gathered from Thompson Financial Datastream, with the time horizon ranging from the first quarter 1950 to the last quarter 2009. As it is standard in related literature, the business cycle is considered as the cyclical component whose recurring movements range from 6 to 32 quarters. This conception will be handy for

[4] All the R (R Development Core Team, 2007) routines developed to implement the procedures detailed in this section are available from the authors upon request.

window length selection, as well as for electing the principal components to be discarded from the cyclical analysis. In effect, given that we are interested in the follow-up of regular dynamics of up to 8 years, setting a window length of 32 quarters becomes the natural choice. This leads us to the panel of principal components which is depicted in Figure 1.[5]

[**Insert Figure 1 about here**]

As mentioned above, the focus over frequencies above 6 and below 32 quarters, will provide guidance for discarding uninteresting principal components for business cycle extraction. Hence, this enables us to dispense with the first two principal components of Figure 1, which are assigned to much lower frequencies than the ones of interest. In particular, PC1 is linked to a slow-moving component (trend), while PC2 is associated to movements with a frequency noticeably larger than 32 quarters. By a similar token, all principal components above the ninth are not considered relevant for the purposes of the current analysis, as they take control of short fluctuations markedly below 6 quarters. Our business cycle indicator is thus composed by summing the remainder (3–9) principal components (henceforth CRR-filter).

[**Insert Figure 2 about here**]

Figure 2 represents the business cycle obtained with the CRR-filter against three other business cycle indicators, to wit: Christiano-Fitzgerald (CF); Baxter-King (BK); and Hodrick-Prescott (HP). Some comments about this figure are in order. First, the aforementioned high-pass features of the HP-filter are visible in its bumpy behavior. Second, both the CF

[5] For the sake of parsimony, only the first 12 PCs are plotted here; the remainder components are associated with very high-frequency movements holding a negligible share of about $0.2 \times 10^{-4}\%$ of total explained variance.

and BK have a similar behavior which is a consequence of their band-pass attributes. However, as mentioned earlier the BK-filter is unable to yield end-of-sample estimates. Lastly, as it it clear from the examination of this figure, the herein proposed indicator possesses a similar behavior to the remainder and hence, in this sense, it can be thought as an alternative method for characterizing the cyclical dynamics of economic activity.

[**Insert Figure 3 about here**]

The behavior of the CRR-filter is also patent in Figure 3, wherein the proposed indicator is plotted against a set of shaded regions representing recession periods, from peak (**P**) to through (**T**), dated by US Business Cycle Dating Committee. As the graphical inspection of this figure puts forward, the business cycle indicator yielded by dint of the method proposed herein is broadly in line with the contractions and expansions dated by the NBER.

3.2 A Real-Time Exercise

From the policymaker stance a feature of remarkable importance is the reliability of real-time estimates of the business cycle indicator (see, for instance, Orphanides and van Norden, 2002). Here and below, by a real-time estimate we mean the business cycle estimate, conditional on the information set available at such point in time. Ex post revision of the estimates is typically due to either published data revision, or to recomputations given the arrival of further data on subsequent quarters. Here, a fixed data set is used so that the unique source of revision is the latter. As advocated by Orphanides and van Norden (2002), recomputations are responsible for a large share of recurrent revision in the estimates.

[**Insert Figure 4 about here**]

In Figure 4 we represent real-time business cycle estimates against final estimates for the CRR-filter, CF-filter and HP-filter. From all the filters used in the preceding subsection, only the BK-filter is not considered for the reliability assessment given that it is not able to produce end-of-sample estimates. The period under consideration ranges from the first quarter 1998 to the last quarter 2009 and corresponds to 20% of the sample size.

[**Insert Table 1 about here**]

In order to provide a portrait of the revision features of the aforementioned business cycle indicators, we report in Table 1 a set of realiability measures. The set of statistics considered here is commonly employed in related literature, encompassing the contemporaneous correlation of real-time against final estimates, the noise-to-signal ratio and the signal concordance. From the inspection of the obtained results it is clear that the real time performance of the HP-filter is in the overall dominated by the CF-filter. This is in line with one's expectations, being also found in other comparisons performed in the literature (see, among others, Azevedo, Koopman and Rua, 2006). From the examination of this table one can also observe that the business cycle indicator proposed in this paper dominates in all the above mentioned measures the CF and HP filters. The correlation of the CRR-filter is almost 0.9 and hence much higher than the ones obtained by the HP and CF filters (0.65 and 0.67, respectively). In addition, the noise-to-signal ratio is quite lower reinforcing the relative information content of the real-time estimates obtained with the CRR-filter. Finally, the share of time that the real-time estimates of the CRR-filter have the same sign as final estimates is higher in comparison with the HP and CF filters.

4. SUMMARY AND CONCLUSIONS

This paper introduces a business cycle indicator based on singular spectrum analysis. The proposed business cycle indicator is shown to bear a resemblance with band-pass filtered output. Given that the end-of-sample behavior is frequently a thorny issue in business cycle assessment, a real-time estimation exercise is here executed in order to evaluate the reliability of the several filters. As discussed earlier, one should bear in mind that it is extremely difficult to estimate output gap in real time, given that only with time one can be more precise about the "true" position at a given moment. Our results suggest that the business cycle indicator proposed herein is endowed with a better revision performance than other filters often applied in the literature.

Table 1: Reliability Statistics Over Different Business Cycle Indicators.

Method	Reliability Statistic		
	Correlation	Noise-to-signal	Sign concordance
Hodrick-Prescott	0.649	0.844	0.583
Christiano-Fitzgerald	0.669	0.750	0.583
Carvalho-Rodrigues-Rua	0.893	0.553	0.646

NOTES: Correlation corresponds to Pearson correlation coefficient between real-time estimates and final estimates of the business cycle. Noise-to-signal corresponds to the ratio of standard deviation of revisions and the standard deviation of final estimates. Sign concordance assesses the percentage of periods under which the sign of real-time estimates and final estimates coincide. All reliability measures are computed for the period of 1998(Q1)–2009(Q4).

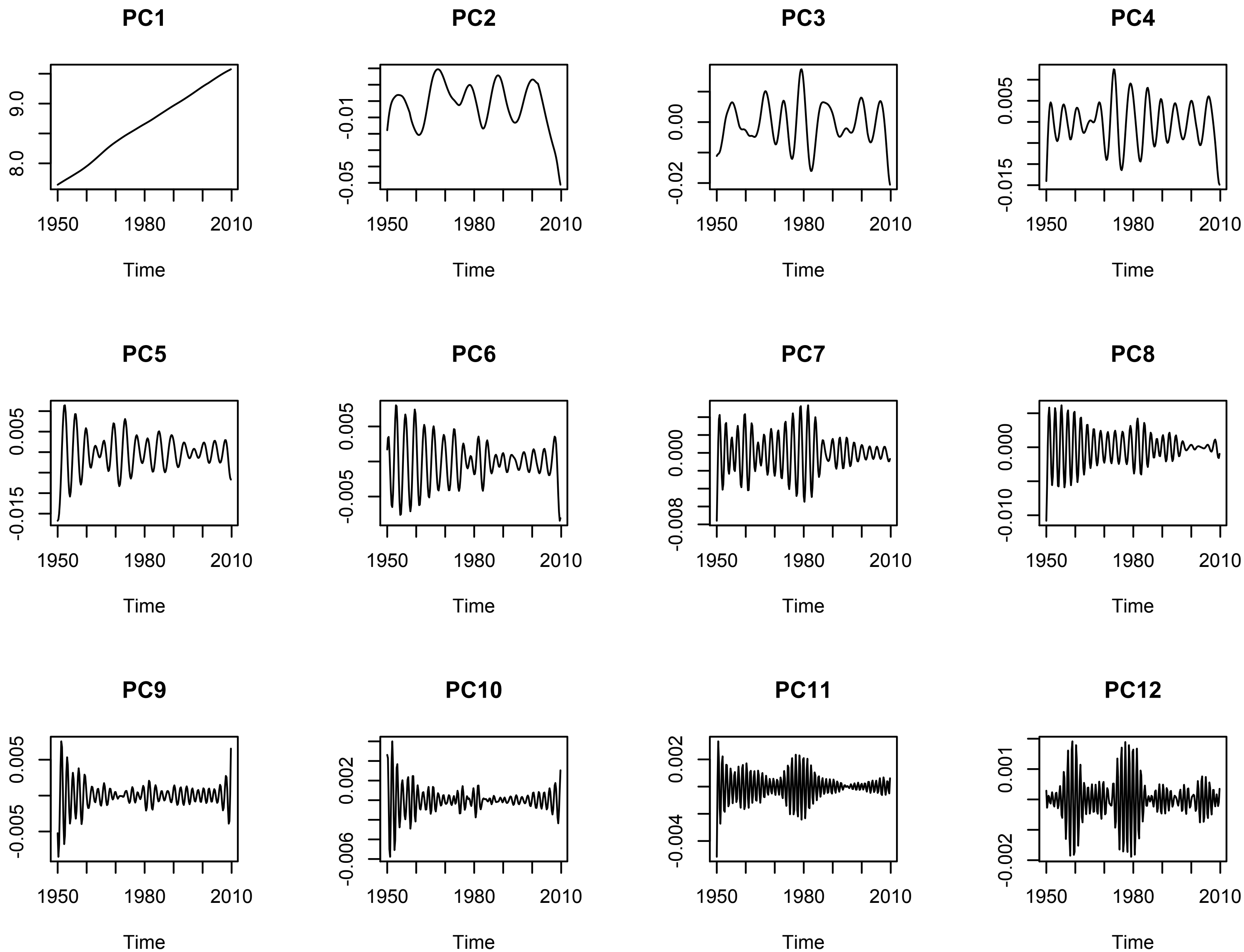


Figure 1: A Panel of the First 12 Principal Components Plotted as Time Series.

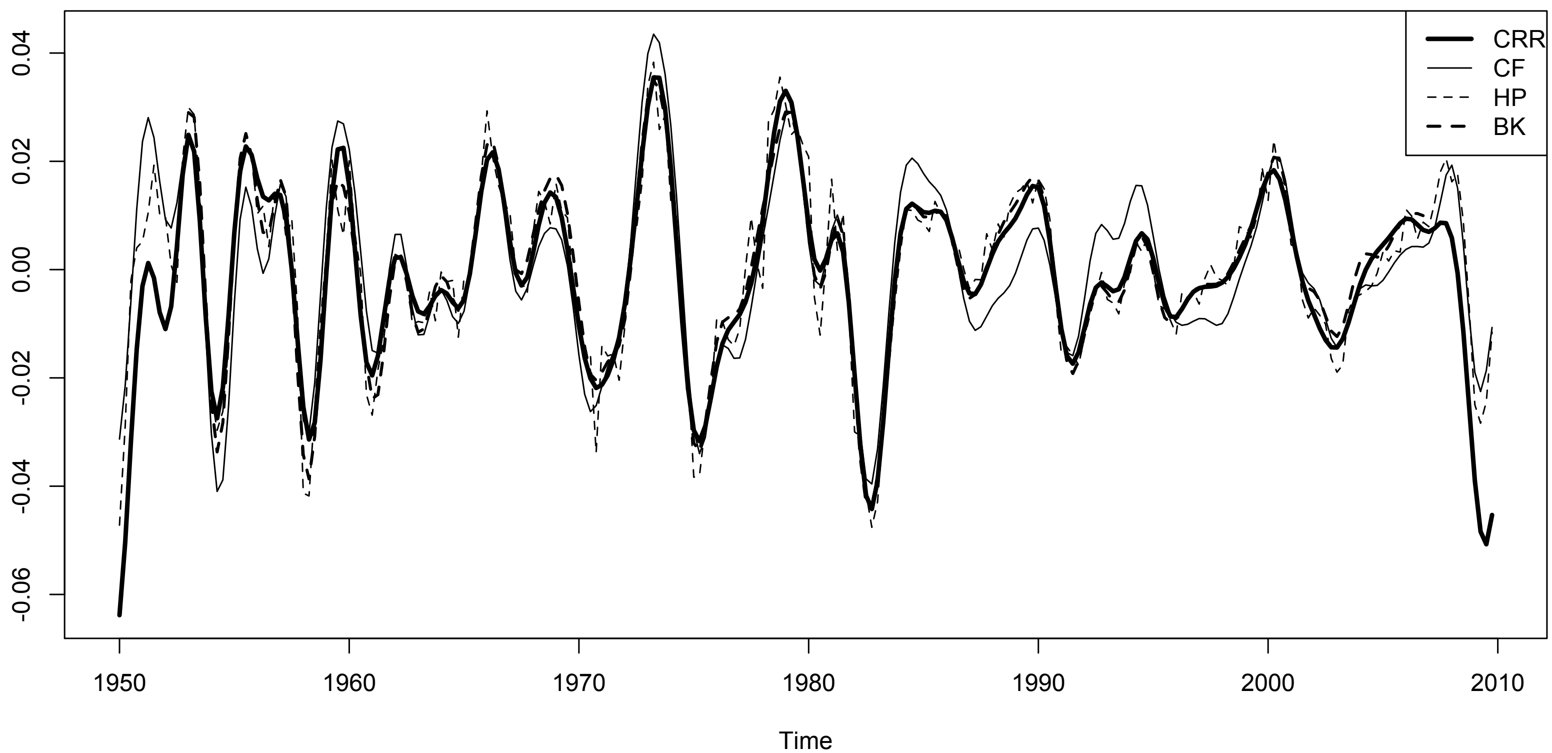


Figure 2: Comparison of the Following Business Cycle Indicators: (CRR) Carvalho-Rodrigues-Rua Filter; (CF) Christiano-Fitzgerald Filter; (BK) Baxter-King; (HP) Hodrick-Prescott Filter.

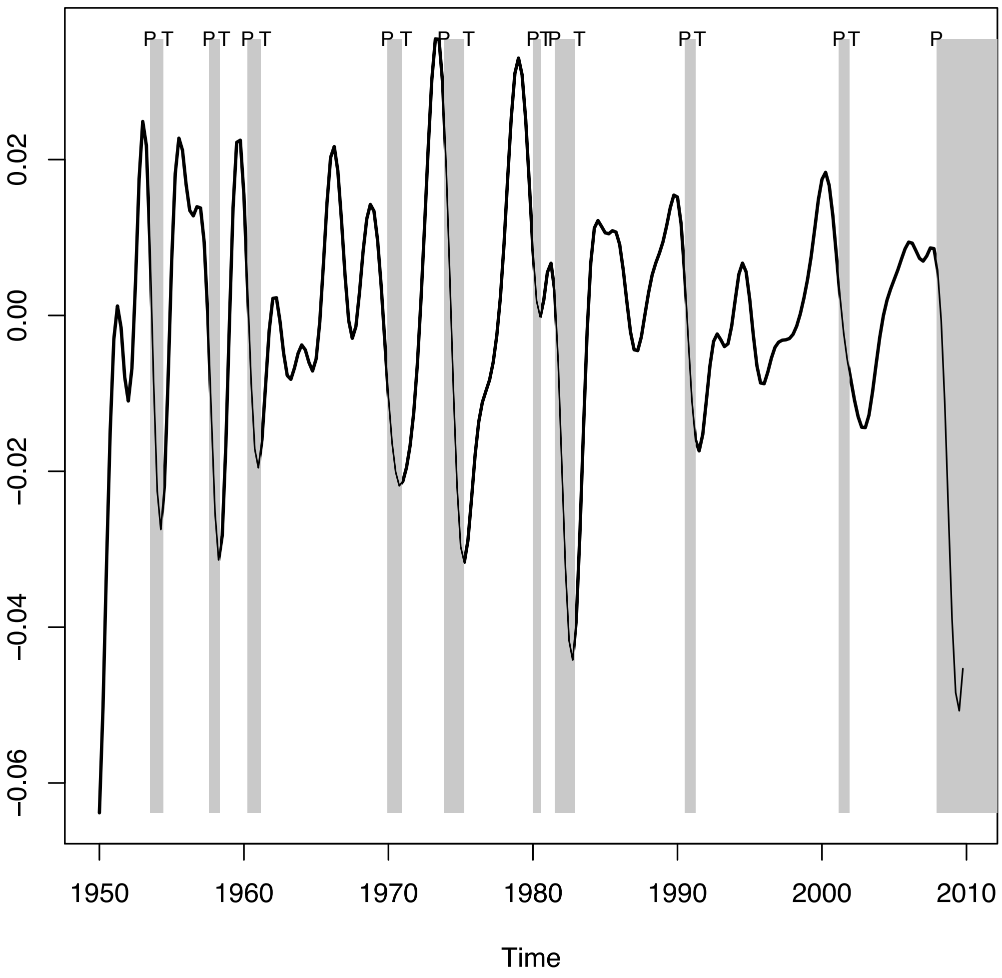
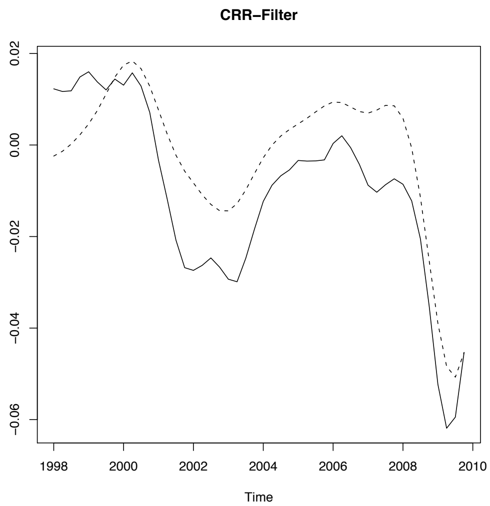
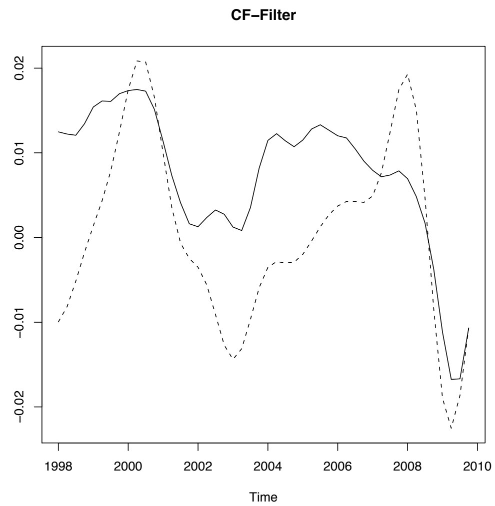


Figure 3: US Business Cycle Indicator Obtained Through the CRR-Filter and the recession periods (shaded areas), from peak (**P**) to through (**T**), dated by the US Business Cycle Dating Committee of the National Bureau of Economic Research (NBER).



(a)



(b)

(c)

Figure 4: Comparison of Final (——) and Real-Time Estimates (- - -) for the Following Business Cycle Indicators: (a) CRR – Carvalho-Rodrigues-Rua; (b) CF – Christiano-Fitzgerald Filter; (c) HP – Hodrick-Prescott Filter. The period under consideration ranges from 1998(Q1)–2009(Q4).

REFERENCES

Allen, M.R. and Smith, L.A. (1996). "Monte Carlo SSA: Detecting Irregular Oscillations in the Presence of Colored Noise," *Journal of Climate*, 9, 3373–3404.

Azevedo, J.V., Koopman S.J. and Rua, A. (2006). "Tracking the Business Cycle of the Euro Area: a Multivariate Model-Based Bandpass Filter," *Journal of Business and Economic Statistics*, 24, 278–290.

Basilevsky A and Hum, D. (1979). "Karhunen Loève Analysis of Historical Time Series with an Application to Plantation Births in Jamaica," *Journal of the American Statistical Association*, 74, 284–290.

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

Broomhead, D.S. and King, G.P. (1986). "Extracting Qualitative Dynamics from Experimental Data," *Physica D*, 20, 217–236.

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

Christiano, L. and Fitzgerald, T. (2003). "The Band-Pass Filter," *International Economic Review*, 44, 435–465.

Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*, Chapman & Hall/CRC, London.

Hassani, H., Heravi, S. and Zhigljavsky, A. (2009). "Forecasting European Industrial Production with Singular Spectrum Analysis," *International Journal of Forecasting*, 25, 103–118.

Hodrick, R.J. and Prescott, EC. (1997). "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29, 1–16.

Kondrashov, D. and Ghil, M. (2006). "Spatio-Temporal Filling of Missing Points in Geophysical Data Sets," *Nonlinear Processes in Geophysics*, 13, 151–159.

Loève, M. (1978). *Probability Theory II*, 4th Edition. Springer Verlag, New York.

Orphanides, A. and van Norden, S. (2002). "The Unreliability of Output-Gap Estimates in Real Time," *The Review of Economics and Statistics*, 84, 569–583.

Paegle, J.N., Byerle, L.A. and Mo, K.C. (2000). "Intraseasonal Modulation of South American Summer Precipitation," *Monthly Weather Review*, 128, 837–850.

Pollock, D.S.G. (2000). "Trend Estimation and De-Trending Via Rational Square-Wave Filters," *Journal of Econometrics*, 99, 317–334.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Ravn, M.O. and Uhlig, H. (2002). "On Adjusting the Hodrick-Prescott Filter For the Frequency of Observations," *The Review of Economics and Statistics*, 84, 371–380.

Vautard, R. and Ghil, M. (1989). "Singular Spectrum Analysis in Nonlinear Dynamics, with Applications to Paleoclimatic Time Series," *Physica D*, 35, 395–424.

Yogo, M. (2008). "Measuring Business Cycles: A Wavelet Analysis of Economic Time Series," *Economics Letters*, 100, 208–212.