1 Introduction

Traditionally, the investigators looking for economic fluctuations or trying to disprove their existence have considered the actual macroeconomic time series as a linear combination of distinct, not necessarily mutually independent components. As a rule, one distinguishes between the following components: (a) trend (if the series is not stationary, displaying secular upward or downward movement); (b) cycle(s), which is one of the most disputable concepts of macroeconomic analysis, duration of which ranges from 2-8 years (common business cycles) to 40-60 years (the so-called long, or Kondratiev, waves); (c) seasonal component (systematic movements within one year); and finally (d) white noise (serially uncorrelated, stationary component). Since all these components are not observable, there has been a large variety of different signal-extraction methods (SETs) proposed. By the very same reason it is virtually impossible to judge, which of these methods uncovers the "true trend" or "true cycle".

Even when the methods depart from the same definition of cycle, they can pretty easily give rise to the completely different pictures of the macroeconomic fluctuations. This has extremely important implications for the formulation and testing of the theoretical models of the business cycle as well as for the making and evaluating economic policy decisions, since it affects the stylized facts (empirical regularities) and turning points dating of the business cycles, which are based on the estimates of the cyclical component of the original series.

Therefore some criteria are needed to choose among these methods those which are the most adequate to the problem of the business cycle research. The principal criteria one can use to "sort" these methods from the standpoint of their usefulness for the analysis of economic fluctuations are the *strictness* of the underlying model (that is, how restrictive are the constraints imposed by the corresponding technique on the trend/cycle behavior), the *degree of the distortions* introduced by each method into the signal(s), and the *ability to predict* the cyclical movements of the economy in the future.

In this paper we concentrate upon the "distortion aspect" to test the signal extracting techniques. One can test the signal-extraction methods looking at them from different angles. Those depend on the uses one makes of the business cycle estimates. The most important use of the estimates of the cyclical components are (1) the determination of the so-called stylized facts (moments of various orders) to build and check the macroeconomic models, (2) localization of the turning points of the business cycle, which may serve as a basis for the forecasting of the cycle and hence of timely decision-making in the area of the countercyclical policy.

However, the problem of utmost importance is that the SETs give different results both in terms of the stylized facts and in terms of the cycle chronology. Therefore, since we do not know the true DGP of the time series observed in reality, we cannot choose among these SETs to say which of them uncover the truth. This ignorance undermines all the conclusions made based on any of the detrending methods. Indeed, how can we say that, for example, our model replicates the reality well, if we do not even imagine this reality?

This paper tries to test several signal-extraction techniques using as a criterion their performance in terms of the stylized facts. The attempts to figure out the impact of the techniques on the stylized facts have been already made in the literature. One can mention, for example, Canova (1995), (1998a), and (1999). However, our research differs from the previous ones in that it applies the signal-extraction techniques to the simulated data with known properties and not to the actual macroeconomic time series, as it does, for instance, Canova (1998a). Thus, we possess the exact knowledge about all the moments of the generated series. In other words, we know the true stylized facts.

The idea is to test various signal-extraction techniques from the viewpoint of the distortions they introduce into the stylized facts of the extracted cycles. Hence it would be natural to proceed as follows:

1. Generate signal (cycle) with known stylized facts, e.g. first and second order moments.

2. Add this signal to some kind of trend, apparently reminding those we observe in the real life.

3. Use a detrending technique to extract the hidden signal.

4. Compute the moments in question of the estimate of the signal and calculate an aggregate measure of deviation of these moments from those of the generated true signal as well as some measures of the similarity of the simulated and extracted series.

5. Do the above exercise for different signal-extraction techniques for a large enough number of draws, average the results over all the draws, and compare the resulting aggregate measures of deviations across these methods.

2 Data-Generating Processes

We simulated two signal series, each of which consists of two components: the common component, which makes them interdependent, and the idiosyncratic component, which introduces their own dynamics.

To these cyclical series some trends were added so to create the trended series of the type we generally observe in the real life. We have considered five principal cases: (1) distinct deterministic trends; (2) common deterministic trend; (3) distinct stochastic trends; (4) common stochastic trend independent of the cyclical components; (5) common stochastic trend correlated with the cyclical components. In each case the cyclical components were generated in the same way. The only difference was the DGP of the trend.

The cyclical components are constructed as follows:

$$c_{1t} = \alpha_1 u_t + \beta_1 \varepsilon_{1t}$$

$$c_{2t} = \alpha_2 u_t + \beta_2 \varepsilon_{2t}$$

$$u_t \sim NID(0,1)$$

$$\varepsilon_{1t} = \rho_1 \varepsilon_{1t-1} + \eta_{1t}$$

$$\varepsilon_{2t} = \rho_2 \varepsilon_{2t-1} + \eta_{2t}$$

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \sim NIID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

The disturbance term u_t is common for both series (c_{1t} and c_{2t}), while ε_{2t} and ε_{3t} are mutually independent — since $E(\eta_{1t}\eta_{2s}) = 0 \ \forall t, s$ — and represent the specific component of each signal series. The common component is not correlated with the specific components at any lag or lead: $E(u_t \eta_{is}) = 0 \ \forall t, s$ and $i = 1, 2$.

The trended series for the cases (1) and (3) were created in the following way:

(1) *Distinct deterministic trends*:

$$y_{1t} = a_1 t + b_1 t^2 + c_{1t}$$

$$y_{2t} = a_2 t + b_2 t^2 + c_{2t}$$

The first two summands on the right-hand side form the trend of the observed series, y_{1t} and y_{2t}, while the last one (c_{1t} or c_{2t}) is the cyclical component, or signal.

The parameters α_1, α_2, β_1, β_2 determine the relative importance of the common and specific shocks in the dynamics of both cyclical components. The autoregressive parameters, ρ_1 and ρ_2, introduce the "long memory" into the specific component processes and, in turn, into the cyclical components themselves. Thus, if ρ_1 and ρ_2 are different from zero, the transitory components, c_{1t} and c_{2t}, are no longer white noise, but display certain persistence in their behavior. We think that this may be a more realistic approach to the cyclical component, than that reducing this component to a mere white noise.

(2) *Common deterministic trend* is obtained from the previous DGP by imposing two constraints: $a_1 = a_2$ and $b_1 = b_2$.

(3) *Distinct stochastic trends*:

$$y_{1t} = g_{1t} + c_{1t}$$

$$y_{2t} = g_{2t} + c_{2t}$$

where the trends g_{1t} and g_{2t} are:

$$g_{1t} = \delta + g_{1t-1} + \xi_{1t}$$

$$g_{2t} = \delta + g_{2t-1} + \xi_{2t}$$

The noise terms, ξ_{1t} and ξ_{2t}, are independently and identically distributed:

$$\begin{pmatrix} \xi_{1t} \\ \xi_{2t} \end{pmatrix} \sim NIID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

(4) *Common stochastic trend independent of the cyclical components* is defined as follows:

$$y_{1t} = g_t + c_{1t}$$

$$y_{2t} = g_t + c_{2t}$$

where

$$g_t = \delta + g_{t-1} + \xi_{2t}$$

(5) *Common stochastic trend correlated with the cyclical components* is defined almost as the process (4), the only difference being:

$$g_t = \delta + g_{t-1} + u_t$$

So that the trend is correlated both with c_{1t} and with c_{2t} via their common disturbance, u_t.

The theoretical moments of the processes c_{1t} and c_{2t}, regardless of the nature of the trends, are computed as follows:

$$E(c_{1t}) = E(c_{2t}) = 0$$

$$Cov(c_{1t}, c_{1s}) = \begin{cases} \alpha_1^2 + \frac{\beta_1^2}{1-\rho_1^2}, & \text{if } t = s \\ \frac{\rho_1^{|t-s|}}{1-\rho_1^2}\beta_1^2, & \text{otherwise} \end{cases}$$

Analogously,

$$Cov(c_{2t}, c_{2s}) = \begin{cases} \alpha_2^2 + \frac{\beta_2^2}{1-\rho_2^2}, & \text{if } t = s \\ \frac{\rho_2^{|t-s|}}{1-\rho_2^2}\beta_2^2, & \text{otherwise} \end{cases}$$

Also,

$$Cov(c_{1t}, c_{2s}) = \begin{cases} \alpha_1\alpha_2, & \text{if } t = s \\ 0, & \text{otherwise} \end{cases}$$

The third order central moment:

$$Skewness(c_{1t}) = Skewness(c_{2t}) = 0$$

The fourth order central moment:

$$Kurtosis(c_{1t}) = Kurtosis(c_{2t}) = 0$$

The length of each simulated series was chosen to be 250 observations — a standard length for the Post World War II macroeconomic variables measured at quarterly frequency. There were 500 independent draws made, so that we got 500 pairs of the trended series for each of the five cases corresponding to different DGPs of the trends.

3 Estimation

We tried the following ten signal-extraction methods: First-order difference (FOD), Linear trend filter (LT), Quadratic filter, or second-order polynomial (QT), Median filter (MED), Hodrick-Prescott filter (HPF), Bandpass filter (BPF), Frequency-domain filter (FDF), Locally weighted regression (LWR), Caterpillar (or "Cat" for shortness), Aoki's method, and West's method. The details of the algorithms and motivation of some of these methods an interested reader can find in: HPF — Hodrick and Prescott (1997); BPF — Baxter and King (1995); LWR — Heiler (1999); MED — Wen and Zeng (1999); Caterpillar — Danilov and Zhigljavsky (1997), Golyandina, Nekrutkin and Zhigljavsky (2001) or Ghil and Yiou (1996), Ghil and Taricco (1997); Aoki's method — Aoki (1994); West's method — West, Prado, and Krystal (1999).

To compare the performance of different signal-extraction techniques we introduced and computed the average measure of deviations (AMD), which measures the average distance between the first four moments (means, variances, autocovariances, covariances, skewnesses, and kurtosises) of the simulated and extracted series:

$$AMD = \frac{1}{m}\sum_{i=1}^{m} |\theta_i^s - \theta_i^e|$$

where θ_i^s is a moment of the simulated series, and θ_i^e is a moment of the extracted signal series.

Another measure is constructed in such a way that it is bounded within $[0, 1]$ interval and that the impact of two large discrepancies between some moments is attenuated. It is called average exponential measure of deviations ($AMDE$) and it is computed as follows:

$$AMDE = \frac{1}{m}\sum_{i=1}^{m} \{1 - \exp(-|\theta_i^s - \theta_i^e|)\}$$

It is also calculated for the first four moments as AMD. In order to disentangle the influence of the second-order moments, which are currently the most demanded by the business cycle researchers, we computed $AMDE$ for these moments only. It is called $AMDE2$ in the below discussion. Both measures are inversely related to the performance: the closer $AMDE$ or $AMDE2$ is to 1, the poorer is the performance of the corresponding SET.

There was also one standard measure of the similarities between two series used, namely the Theil's inequality coefficient (Theil). The discussion of this measure can be found, for instance, in Pindyck and Rubinfeld (1991, pp. 337-341). The Theil's inequality, U, is defined as:

$$U = \frac{\sqrt{\sum_{i=1}^{m}(\theta_i^s - \theta_i^e)^2}}{\sqrt{\sum_{i=1}^{m}(\theta_i^s)^2} + \sqrt{\sum_{i=1}^{m}(\theta_i^e)^2}} \tag{1}$$

The results of the simulation are summarized in the tables of the Appendix. Tables 1a-1c contain the true moments and the sample moments of the simulated and extracted cyclical series. In the Tables 2a-2e the aggregate measures of the dissimilarities are presented.

Tables from 1a to 1c compare the true moments with the sample moments of extracted series across all the ten signal-extraction techniques considered. Since the results for the processes with distinct and common deterministic trends are similar as well as those for the processes with distinct and common stochastic trends not correlated with the cycles, in the Appendix we reproduce only three tables. The tables with the sample moments for the processes with common deterministic trend and common stochastic trend not correlated with cycles are available on request.

In order to judge how significant are the differences between the true moments and the moments of the extracted series, we have found the 1st and the 99th percentiles of the empirical distributions of the moments of the estimates of the signal. If the true moment lies outside of the interval between the 1st and the 99th percentiles, then the difference between the true and estimated moment is thought to be significant. In order to save space, we do not report the percentiles. The results of the computation of the extreme percentiles conform to those of calculation of the mean.

Under the deterministic trend specification (see Table 1a) most of the methods, save FOD and LWR, estimate well the means of the cycles. FOD and LWR extract cycles with the mean equal to the slope of the deterministic cycle. However, most of the SETs, except for QT and LT, underestimate the variances and covariances of each individual series. In the same time QT, LT, HPF, BPF, and Caterpillar give quite good estimates of the crosscovariances. FOD overestimates the crosscovariance at lag 0 and makes the two cycles dependent at nonzero lags. West's method, FDF, MED, and LWR underestimate the crosscovariance of x_1 and x_2 at zero lag. The higher moments seem to be not distorted by any of the SETs, with the exception of MED and LWR, which extract cycles with very high kurtosises, i.e. with large tail distributions.

Under the stochastic trend uncorrelated with cycles assumption (see Table 1b) the means and skewnesses remain undistorted by filtering. The largest distortions are introduced into the second-order moments, especially those for each individual series. The covariances and autocovariances of the cycles estimated by QT, LT, HPF, BPF, Caterpillar, and West's technique are significantly overestimated. FOD overestimates only the variance and renders autocovariance at higher lags negative, making the cycle remind an MA(1) process with negative coefficient. FDF and MED still underestimate the second-order moments. HPF, BPF, and Caterpillar reproduce the crosscovariances quite closely. It seems that the switch from the deterministic to the stochastic trends, which are not correlated to the cycles, does not affect for these methods the relationships between the cyclical components. Under the stochastic trends more SETs (namely QT, LT, West's method, MED, and LWR) are distorting the fourth-order moments of the distribution of the cycle.

The assumption of stochastic trend correlated with the cycles (see Table 1c) leads to the drastic changes, which affect mainly the second-order moments. First, it introduces a kind of asymmetry between the cyclical components: the one, which is positively correlated with trend, has much greater variance than the cycle which is correlated negatively. It appears that positive (negative) correlation between the trend and cycle accentuates (attenuates) the variability of the observable series and hence the SETs extract cycles with higher (lower) variance than in the case with no correlation. The crosscovariance at lag zero changes its sign (now it is positive) and gets much higher absolute value. Moreover, the DGP specification change induces a kind of temporal asymmetry: the estimates of cycles are correlated higher at negative lags than at the positive ones. In other words, cycle 1 starts to lead slightly cycle 2. The above observations are true for all SETs, save West and FDF. The first- and third-order moments remain as usual intact. The kurtosises are significantly negative for QT and LT and positive for West, MED, and LWR.

The fact that under deterministic trend the second-order moments are generally underestimated, while under the stochastic trend they are overestimated is because the trend extracted by many of the SETs is not smooth enough compared to the deterministic one and is too smooth compared to the stochastic trend. Therefore in case of deterministic trend, the SETs do not extract all the cyclical component, adding part of it to the trend, while in the case of stochastic trend these methods include into the estimate of the cycle part of the trend variability. This "leakage" of variability between different time series components explains also (in the case of the common trend correlated with the cycles) the positive crosscovariances between the estimates of the cycles when the true cycles are negatively correlated.

These results can be formalized and summarized as follows. Under any DGP the estimate of the cyclical component is some linear combination (if, of course, we are using a linear filter as our SET) of the components of the original time series. If we assume the series to be composed only of trend and cycle, then the estimate of the cycle, $\hat{c}_{it}$, would be:

$$\hat{c}_{it} = f(y_{it}) = f(g_{it} + c_{it}) = f(g_{it}) + f(c_{it})$$

where $f(\cdot)$ is the impulse-response function of a SET; y_{it} is the original time series; g_{it} is a true trend; c_{it} is a true cycle.

Since the objective of a SET extracting cyclical component is to eliminate non-stationarity from the original series, its impulse-response function can be represented as follows:

$$f(L) = \sum_{j=0}^{\infty} \alpha_j L^j$$

where L is the lag operator; α_j are the SET-specific weights.

In the case of deterministic trends the original series is defined as follows

$$y_{it} = h_i(t) + c_{it}$$

where $h_i(t)$ is some deterministic function of time, for example, an r-th order polynomial.

The cycle can be expressed as result of interactions of common and idiosyncratic shocks:

$$c_{it} = \omega_{i1} u_t + \omega_{i2} \sum_{j=0}^{\infty} \theta_{ij} L^j \varepsilon_{it}$$

where ω_{ik} is some number (not necessarily positive) determining relative weights of common and idiosyncratic components of cycle as well as degree and sign of correlation between the cycles; u_t is common shock, ε_{it} is idiosyncratic shock.

The variance of the true cycle is

$$Var(c_{it}) = \omega_{i1}^2 \sigma_u^2 + \omega_{i2}^2 \left(\sum_{j=0}^{\infty} \theta_{ij}^2 \right) \sigma_{\varepsilon i}^2$$

and the covariance between two true cycles would be

$$Cov(c_{1t}, c_{2t}) = \omega_{11} \omega_{21} \sigma_u^2$$

This is true for any specification of the trend: both deterministic and stochastic. However, the moments of the extracted cycle, $\hat{c}_{it}$, will be different, depending on the DGP of the trend because of the "leakage" of variability between the estimated components produced by the SETs.

Deterministic trend(s):

$$\hat{c}_{it} = \sum_{j=0}^{n} \alpha_{ij} L^j \left\{ h_i(t) + c_{it} \right\} = \sum_{j=0}^{n} \alpha_{ij} L^j \left\{ h_i(t) + \omega_{i1} u_t + \omega_{i2} \sum_{j=0}^{\infty} \theta_{ij} L^j \varepsilon_{it} \right\}$$

More compactly this can be expressed as

$$\hat{c}_{it} = \tilde{h}_i(t) + \omega_{i1} \sum_{j=0}^{n} \alpha_{ij} L^j u_t + \omega_{i2} \sum_{j=0}^{\infty} \tilde{\theta}_{ij} L^j \varepsilon_{it}$$

where $\tilde{h}_i(t) = (1-L)^d \sum_{j=0}^{\infty} \alpha_{ij} L^j h_i(t)$, $\sum_{j=0}^{\infty} \tilde{\theta}_{ij} L^j = \sum_{j=0}^{n} \alpha_{ij} L^j (\sum_{j=0}^{\infty} \theta_{ij} L^j)$; d is the order of differencing. We make impulse-response function of a SET dependent of the series being detrended, since some of the methods are data-adaptive and hence their weights are influenced by the time series to which these filters are applied.

Hence the variance of the estimate of cycle is

$$Var(\hat{c}_{it}) = \omega_{i1}^2 \left(\sum_{j=0}^{n} \alpha_{ij}^2 \right) \sigma_u^2 + \omega_{i2}^2 \left(\sum_{j=0}^{\infty} \tilde{\theta}_{ij}^2 \right) \sigma_{\varepsilon i}^2$$

For example, for FOD filter, whose impulse-response function is $f^{FOD}(L) = 1 - L$, the variance of the extracted cycle will be $Var(\hat{c}_{it}) = E(c_{it} - c_{it-1}) = 2\gamma_{ci}(0) - 2\gamma_{ci}(1)$, where $\gamma_{ci}(\tau)$ is autocovariance of c_{it} at lag τ. For the cycle's specification we are using here this implies that $Var(\hat{c}_{it}) = 2 * (0.944 - 0.556) = 0.776$, which can be confirmed upon consulting Table 1a of Appendix, where the variances of the cyclical estimates oscillate around 0.77. For HPF, BPF and Caterpillar the variance is even smaller, since they have $\sum_{j=0}^{n} \alpha_{ij}^2 \in (0,1)$ and $\sum_{j=0}^{\infty} \tilde{\theta}_{ij}^2 < \sum_{j=0}^{\infty} \theta_{ij}^2 = \frac{1}{1-\rho^2} = 2.77$ (recall that we generated idiosyncratic part of the cycle as an AR(1) with autoregressive coefficient $\rho = 0.8$). For instance, for BPF the former varies from 0.12 to 0.19, while the latter is approximately equal to 1.75, depending on the number of terms in the MA representation of the filter's impulse-response function, n.

Similarly, the cross-covariance is

$$Cov(\hat{c}_{1t}, \hat{c}_{2t}) = \omega_{11}\omega_{21} \left(\sum_{j=0}^{n} \alpha_{1j}\alpha_{2j} \right) \sigma_u^2$$

For instance, when we are employing a FOD, the cross-covariance of the extracted cycle will be twice as large as the covariance between the true cycles:

$$Cov(\hat{c}_{1t}^{FOD}, \hat{c}_{2t}^{FOD}) = 2\left[\gamma_{c12}(0) - \gamma_{c12}(1)\right] = 2\gamma_{c12}(0)$$

because, fortunately for us, $\gamma_{c12}(1) = 0$.

We can see that this is the case looking at the cross-covariance at lag zero of the FOD-extracted cycles in Table 1a.

The picture becomes more complicated when we turn to the case of stochastic trends. Let's use for the sake of simplicity the $I(1)$ trend, that is, $g_{it} = g_{it-1} + v_{it}$. The trend shocks, v_{1t} and v_{2t}, are assumed to be independently and identically distributed and mutually uncorrelated as well as not correlated to the cyclical components, c_{1t} and c_{2t}. Then the estimate of cyclical component will look like follows:

$$\hat{c}_{it} = \sum_{j=0}^{\infty} \alpha_{ij} L^j \left\{ (1-L)^{-1} v_{it} + \omega_{i1} u_t + \omega_{i2} \sum_{j=0}^{\infty} \theta_{ij} L^j \varepsilon_{it} \right\}$$

which can be expressed also as

$$\hat{c}_{it} = \sum_{j=0}^{n} \tilde{\alpha}_{ij} L^j v_{it} + \omega_{i1} \sum_{j=0}^{n} \alpha_{ij} L^j u_t + \omega_{i2} \sum_{j=0}^{\infty} \tilde{\theta}_{ij} L^j \varepsilon_{it}$$

where $\sum_{j=0}^{n} \tilde{\alpha}_{ij} L^j = (1-L)^{-1} \sum_{j=0}^{\infty} \alpha_{ij} L^j$.

Distinct stochastic trends not correlated with cycles:

$$Var(\hat{c}_{it}) = \left(\sum_{j=0}^{n} \tilde{\alpha}_{ij}^2 \right) \sigma_{vi}^2 + \omega_{i1}^2 \left(\sum_{j=0}^{n} \alpha_{ij}^2 \right) \sigma_u^2 + \omega_{i2}^2 \left(\sum_{j=0}^{\infty} \tilde{\theta}_{ij}^2 \right) \sigma_{\varepsilon i}^2$$

and

$$Cov(\hat{c}_{1t}, \hat{c}_{2t}) = \omega_{11}\omega_{21} \left(\sum_{j=0}^{n} \alpha_{1j}\alpha_{2j} \right) \sigma_u^2$$

In this case the variances of the extracted cycles would be always higher than the variances of the estimates of the cycles obtained in case of deterministic trends, and most probably will be higher than true variances. What for covariances, since the trends' disturbances are not correlated, there will be no difference between the covariances corresponding to the deterministic trends DGP and those corresponding to the distinct stochastic trends.

Common stochastic trends not correlated with cycles:

$$Var(\hat{c}_{it}) = \left(\sum_{j=0}^{n} \tilde{\alpha}_{ij}^2\right) \sigma_v^2 + \omega_{i1}^2 \left(\sum_{j=0}^{n} \alpha_{ij}^2\right) \sigma_u^2 + \omega_{i2}^2 \left(\sum_{j=0}^{\infty} \tilde{\theta}_{ij}^2\right) \sigma_{\varepsilon i}^2$$

and

$$Cov(\hat{c}_{1t}, \hat{c}_{2t}) = \left(\sum_{j=0}^{\infty} \tilde{\alpha}_{1j}\tilde{\alpha}_{2j}\right) \sigma_v^2 + \omega_{11}\omega_{21} \left(\sum_{j=0}^{n} \alpha_{1j}\alpha_{2j}\right) \sigma_u^2$$

When common trend is introduced, not only variances — as in case of distinct stochastic trends — become upward biased, but also the covariances. It is easy to show that for FOD $Var(\hat{c}_{it}^{FOD}) = \sigma_v^2 + 2\gamma_{ci}(0) - 2\gamma_{ci}(1)$, which given that $\sigma_v^2 = 1$, is equal to 1.776. The covariance between two cyclical components at zero lag is: $Cov(\hat{c}_{1t}^{FOD}, \hat{c}_{2t}^{FOD}) = \sigma_v^2 + 2\gamma_{c12}(0)$. Hence $Cov(\hat{c}_{1t}^{FOD}, \hat{c}_{2t}^{FOD}) = 1 - 0.5 = 0.5$. Compare these results with those presented in Table 1b.

Common stochastic trends correlated with cycles:

$$Var(\hat{c}_{it}) = \left\{\left(\sum_{j=0}^{n} \tilde{\alpha}_{ij}^2\right) + 2\omega_{i1} \left(\sum_{j=0}^{\infty} \tilde{\alpha}_{ij}\alpha_{ij}\right) + \omega_{i1}^2 \left(\sum_{j=0}^{n} \alpha_{ij}^2\right)\right\} \sigma_u^2 + \omega_{i2}^2 \left(\sum_{j=0}^{\infty} \tilde{\theta}_{ij}^2\right) \sigma_{\varepsilon i}^2$$

Here we can see the source of the apparent ”asymmetry” between both cyclical components. If the cycle is positively (negatively) correlated with the trend, its variance would be higher (lower) than in the case of distinct stochastic trends. Now, if there is no correlation between the trend and cycles, we will have the extracted cycles' covariance depending only on the variance of the trend disturbance and the covariance between the true cycles.

$$
\begin{aligned}
Cov(\hat{c}_{1t}, \hat{c}_{2t}) \quad = \quad & \left\{ \left(\sum_{j=0}^{\infty} \tilde{\alpha}_{1j} \tilde{\alpha}_{2j} \right) + \omega_{11}\omega_{21} \left(\sum_{j=0}^{n} \alpha_{1j}\alpha_{2j} \right) + \right. \\
& \left. + \omega_{11} \left(\sum_{j=0}^{\infty} \alpha_{1j} \tilde{\alpha}_{2j} \right) + \omega_{12} \left(\sum_{j=0}^{\infty} \tilde{\alpha}_{1j} \alpha_{2j} \right) \right\} \sigma_v^2
\end{aligned}
$$

Here the covariances between trend disturbances, on one side, and the cyclical components, on the other side, may augment or diminish the bias, depending on their signs. In this particular case $\omega_{11} = -\omega_{12}$, and for the non-data-adaptive methods like FOD, HPF, BPF, whose impulse-responses do not depend on the filtered series, the last two terms in the above equation offset each other and hence there is no differences between the covariances under common trend correlated or not with the cycles.

The above discussion implies that the specification of trend's DGP plays crucial role with respect to the stylized facts distortions. The randomness of the trend is itself an important source of bias. The greater is variance of the trend disturbances, the larger will be this bias. For example, for $\sigma_v^2 = 1$ the contribution of the trend disturbances into the variability of the BPF-extracted cycle (with two distinct stochastic trends) is around 1.7, or 78%. If the variance of trend disturbances were smaller, the variance of the extracted cycle would decrease too. Indeed, when we diminish σ_v^2 to 0.5, the variance of the estimated cycle falls up to 0.97, which is pretty close to the true value. However, when σ_v^2 is set to zero, the variance of the extracted cycle is 0.59 and its 99%-th percentile is equal to 0.77, that is, the variance is downward biased.

The last effect is due purely to the filter-dependent distortions. Therefore we have two major causes of bias: specification of trend (stochastic and/or common trend) and the impact of the filter. The first is beyond our control, while the second can be influenced by researcher.

In the case of nonlinear SETs (e.g., MED) it is much more difficult to predict the behavior of the moments of the extracted cyclical components.

Finally, we analyze the aggregate measures of the performance of each signal-extraction technique presented in Tables 2a to 2e. There the SETs are ranked according to the corresponding values of each measure. The higher is the position of a SET in the ranking, the better is its performance.

In what concerns the aggregate measures of performance, the results are pretty similar within each major class of the DGPs. In other words, the ranking and the values of the performance measures are more or less the same both for the processes with distinct and common deterministic trends, on one hand, and for both the processes with distinct and common (correlated or not with the cycles) trends. However, there are big differences between these two groups of the DGPs. In fact, the ranking of the method changes significantly depending on whether one uses deterministic or stochastic trends.

Out of the three performance measures the more similar among them are the ADM and Theil. The latter gives the most stable ranking. The only drastic change in this ranking is the fall of two SETs (QT and LT) from the highest positions under the deterministic trends to the lowest positions under the stochastic trends. This is not surprising, since QT and LT have the same DGP as the deterministic trends. All other methods almost do not change their positions in the ranking. Thus, if we exclude QT and LT, we will see that BPF stays at 1st or 2nd positions, HPF — 1st or 2nd too, Caterpillar — 2nd and 3rd, MED — 1st, 3rd and 4th, FOD — 5th, LWR — 6th,West — 7th and 8th, and FDF — 8th and 10th.

Some idea of the hierarchy and grouping of the SETs can be obtained from the Figures 1 and 2. There the performance measures (ADME and Theil coefficient) of several of the detrending techniques, we are considering here, are plotted against each other. Each point estimate — mean value of the coefficient — is circled into an ellipse whose radii equal to one standard deviation of each measure. Namely

$$x = E(ADME_i) + Var(ADME_i)^{0.5}\cos(\omega)$$

$$y = E(UT_i) + Var(UT_i)^{0.5}\sin(\omega)$$

Although, strictly speaking, this is not an appropriate confidence interval, since the true distribution of the measures is not normal, still serves for the illustration purposes. Thus, one can see that for the deterministic trends there exist two groups of SETs: relatively good performers (HPF, BPF, MixFdw) and bad performers (FOD and MED). When specification is changed to the stochastic trends, the good performers (the composition of the group remains the same) shift northeastward, i.e. their performance worsens and quite significantly — almost twice. Of course, the bad performers move even farther towards the (1,1) point, but now this group is comprised by LT and QT.

4 Improving performance of signal-extraction techniques

The results obtained above show that the performance of all the SETs are not very impressing when the trends are stochastic. On the other hand, there exists certain ranking of SETs, although not perfectly robust to various performance measures. The question is whether we can hope to improve the performance and how to achieve this. It seems that there are at least three ways of improving. The first one is to combine different filters hoping that their linear combination (mixed filter) will do better than its components. The second way is to apply some multivariate filter, which uses more information and is more efficient. The third way would be to use jointly both these approaches, i.e., construct mixed multivariate filters.

The outcomes of constructing mixed and multivariate filters are summarized in Tables 2a-2e of Appendix. There we used the following notation: mixed filter with differentiated weights (MixFdw), mixed filter with equal weights (MixFew), multivariate HPF (MultHPF), mixed multivariate filter with differentiated weights (MMFdw), mixed multivariate filter with equal weights (MMFew).

Mixed filter is just a linear combination of individual SETs. The problem is how to set the weights for this filter. One solution is to use the equal weights: $w_i = \frac{1}{m}$ for $\forall i \in \{1, 2, ..., m\}$, where m is the number of filters used to construct the mixed SET. Another solution can be differentiated weights, depending, for example, on the performance of the individual detrending techniques. For instance,

$$w_i = \frac{UT_i^{-2}}{\sum_{j=1}^{m} UT_j^{-2}}$$

where UT_i is the Theil's coefficient for the $i - th$ SET.

We constructed mixed filter with equal and differentiated weights computed as indicated above. The filters we used in the construction are: MED, FOD, and HPF. They distort the stylized facts least compared with other SET when the trends are stochastic.

It turned out that mixing various filters improves upon the performance of each of them applied separately. The mixed SET performs better even than the best of the filters of which it is comprised. Most of the improvement is due to the combination effect. Making weights dependent on the performance of the individual filters diminishes the stylized facts distortions even more, although not significantly, compared with the equal weights mixed filter. This is because the Theil's coefficients used as a basis for the weights are not very different, thus making the weights themselves quite similar. The above is not true, however, for the deterministic trends case. The mixed filter performs better than FOD and MED, but worse than HPF. The reason is obviously that under this DGP the first two detrending techniques display the worst performance. Therefore by combining it with a better filter we do not gain anything. Moreover, they diverge from the true moments in the same direction — they all underestimate, for instance, the second-order moments and hence do not offset bias introduced by their counterparts included into the mixed filter. The conclusion is that one is likely to improve the performance by constructing the mixed filters out of the SETs producing the cycle estimates, whose sample moments deviate in the opposite directions from the true moments.

Next approach is that of applying *multivariate filters*. One such filter — possibly the simplest one of all the multivariate filters — was proposed in Kozicki (1999). The idea is that in some cases we can suspect the existence of single trend, i.e., the fact that the original time series in question are cointegrated. Therefore we can take advantage of the common trend assumption to detrend these series. One way to do this is to detrend each series separately with a univariate trend and then find the common trend as a linear combination of the estimates of individual trends. The estimates of cycles are computed by subtracting this estimated common trend from each of the observed series.

Multivariate HPF à la Kozicki performs slightly better than its univariate counterpart. Only for the deterministic trends case, when Theil coefficient is used as a performance measure, it seems that multivariate filter is significantly better than the univariate one. However, a drastic improvement is achieved only when the single trend restriction is coupled with mixing various SETs. This leads us to the third solution — that of the *mixed multivariate filter*.

The algorithm may be the following:

1. We get estimates of the trends using several univariate SETs.

2. Mixed filter is constructed using equal or differentiated weights.

3. The individual trend estimates for different series are combined into a single trend. Subtracting this trend from the original time series we receive the mixed multivariate filter estimates of cycles.

As Tables 2d and 2e show, the mixed multivariate filters perform the best. From Table 2b one can see that the mixed multivariate filters go immediately after the SETs, whose structure corresponds to the DGP of the simulated series. Both in the case of single deterministic and stochastic trends the mixed multivariate SET outperforms the techniques used in its composition. Of course, their application makes sense only if we assume the existence of common trend. Again, the difference between the MMFdw and MMFew is not substantial due to little differences between the weights.

5 Summary

In this paper we have compared the performance of various — well known in macroeconometric analysis and relatively novel — signal-extraction techniques in terms of the distortions they introduce into the stylized facts of the simulated cycle. There were 500 pairs series of length 250 observations generated, each of which is constructed as a sum of two components: cycle and trend. Then the signal-extraction techniques were employed to estimate the cyclical component. The business cycle stylized facts (first-, second-, third-, and fourth-order moments) were computed for these cyclical component estimates — dropping the first 25 and the last 25 observations, so that the moments are found for 200 middle observations — and then compared to the sample moments of the simulated series.

There were three aggregate measures of the deviations from the true values calculated. The first three — AMD, AMDE, and AMDE2 are based on the comparison of the means and covariances of the simulated and extracted cycles. The fourth measure — Theil's inequality coefficient directly measures the similarities between the two time series: simulated and extracted.

Under deterministic trends assumption the SETs are usually underestimating the variances and autocovariances of the cycles, while under the stochastic trend DGP they tend to overestimate these moments. Moreover, introduction of the common trend correlated with the cycles leads to the important distortions in the crosscovariances, which were almost exempt of distortions under all other specifications. This can be explained — at least for the linear SETs — by the fact that the signal-extraction methods estimate the cyclical components as a linear combination of true trends and cycles, being unable to separate one from another exactly.

The best SETs, regardless of the DGP used to simulate the time series, in this collection of signal-extraction methods turned out to be BPF, HPF, and Caterpillar. They also give relatively similar results.

One can improve the performance of the SETs by constructing mixed filters, multivariate filters, and mixed multivariate filters. They are better both in terms of ADM and Theil's coefficient. Multivariate SETs, however, deliver rather modest reduction in the distortions. The largest part of the improvement is achieved by mixing filters, that is, by computing their linear combinations. Quite surprisingly, it turns out that linear combination of individual SETs performs better than any of its components. This is a very valuable property, since it allows to minimize the stylized facts distortions.

References

Aoki M. (1994) "Two complementary representations of multiple time series in state-space innovation forms" *Journal of Forecasting* **13**, Issue 2.

Baxter M., King R. (1995) "Measuring business cycles. Approximate band-pass filters for economic time series" *NBER Working Paper # 5022*.

Canova F. (1995) "Does detrending matter for determination of the reference cycle and the selection of turning points?" *Universitat Pompeu Fabra Economics Working Paper Series 113*.

Canova F. (1998a) "Detrending and business cycle facts" *Journal of Monetary Economics* **41**.

Canova F. (1998b) "Detrending and business cycle facts: A user's guide" *Journal of Monetary Economics* **41**.

Canova F. (1999) "Does detrending matter for the determination of the reference cycle and the selection of turning points?" *The Economic Journal* **109**.

Danilov D.L., Zhigljavsky A.A. (eds.). (1997) *Principal Components of Time Series: the Caterpillar Method* (in Russian). University of St.Petersburg Press.

Dorfman J.H., Havenner A. (1997) "An Introduction to model specification and estimation of balanced representation state space models". *A Primer for Users of SSATS 2.0*.

Ghil M., Taricco C. (1997) "Advanced spectral analysis methods". In G.Cini Castagnoli and A.Provenzale (eds.) *Past and Present Variability of the Solar-*

Terrestrial System: Measurement, Data Analysis and Theoretical Models. Bologna: Società Italiana di Fisica (in press).

Ghil M., Yiou P. (1996) "Spectral methods: what they can and cannot do for climatic time series". in D.Anderson and J.Willebrand (eds.) *Decade Climate Variability: Dynamics and Predictability*, Elsevier.

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

Heiler S. (1999) "A Survey of nonparametric regressions" *Universität Konstanz, Zentrum für Finanzen und Ökonometrie Discussion paper*.

Hodrick R.J., Prescott E.C. (1997) "Postwar U.S. business cycles: An empirical investigation" *Journal of Money, Credit, and Banking* **29**.

10. Kozicki S. (1999) "Multivariate detrending under common trend restrictions: Implications for business cycle Research" *Journal of Economic Dynamics and Control* **23**.

Pindyck R.S., Rubinfeld D.L. (1991) *Econometric Models and Economic Forecasting*. NY: McGraw-Hill.

Wen Yi, Bing Zeng (1999) "A simple nonlinear filter for economic time series analysis" *Economics Letters* **64**.

West M., Prado R., Krystal A.D. (1999) "Evaluation and comparison of EEG traces: Latent structure in nonstationary time series" *Journal of American Statistical Association* (to appear).

6 Appendix

The simulation was done using the following values of the parameters:

Cyclical components:

$\alpha_1 = -\alpha_2 = \beta_1 = \beta_2 = 0.5$

$\rho_1 = \rho_2 = 0.8$

(1) Distinct deterministic trends:

$a_1 = 0.5$, $a_2 = 0.3$

$b_1 = -0.0001$, $b_2 = 0.00001$

(2) Common deterministic trend:

$a = 0.5$

$b = -0.0001$

(3) Distinct stochastic trends and (4) common stochastic trend:

$\delta = 0.5$

In the tables we use the following notation: $\mu_i = E(x_{it})$, $\gamma_i(\tau) = cov(x_{it}x_{it-\tau})$, $\gamma_{12}(\tau) = cov(x_{1t}x_{2t-\tau})$, $sk_i = skewness(x_{it})$, and $kurt_i = kurtosis(x_{it})$.

The below statistics were computed for the sample size of 200 observations (250 observations generated originally, dropping then the extreme 50 observations: 25 from the beginning of the sample and 25 from the end of it) for 500 draws.

MixF means "Mixed filter", which is constructed as a linear combination of three SETs: FOD, MED, and HPF. Abbreviation "ew" means "equal weights", while "dw" is "differentiated weights". The latter are calculated based on the Theil coefficients corresponding to each of these techniques when being applied individually.

MMF stands for the "Mixed multivariate filter". It is a mixed filter, which is made also "multivariate" by imposing single trend computed as an arithmetic average of two individual trends.

Table 1a. True cycles' moments and the moments of the extracted cycles

DGP: Distinct deterministic trends

	True	QT	LT	HPF	BPF	Cat	FOD	West	FDF	MED	LWR
μ_1	0	0.0	0.0	0.0	0.0	0.0	0.48	0.0	0.0	0.0	0.48
μ_2	0	0.0	0.0	0.0	0.0	0.0	0.30	0.0	0.0	0.0	0.30
$\gamma_1(0)$	0.94	0.85	0.95	0.59	0.59	0.57	0.77	0.08	0.09	0.26	0.11
$\gamma_1(1)$	0.56	0.46	0.56	0.20	0.19	0.19	-0.27	0.02	0.08	-0.04	0.06
$\gamma_1(2)$	0.44	0.34	0.45	0.10	0.09	0.09	-0.02	-0.02	0.04	-0.04	0.01
$\gamma_2(0)$	0.94	0.86	0.87	0.59	0.58	0.57	0.77	0.08	0.09	0.36	0.11
$\gamma_2(1)$	0.56	0.47	0.48	0.20	0.19	0.19	-0.28	0.02	0.08	0.0	0.06
$\gamma_2(2)$	0.44	0.36	0.37	0.10	0.09	0.09	-0.02	-0.02	0.04	-0.04	0.01
$\gamma_{12}(-2)$	0	0.0	0.0	0.01	0.02	0.01	0.0	0.02	0.0	0.02	0.01
$\gamma_{12}(-1)$	0	0.0	0.0	0.01	0.02	0.02	0.24	0.0	0.0	0.05	-0.01
$\gamma_{12}(0)$	-0.25	-0.25	-0.26	-0.24	-0.23	-0.23	-0.50	-0.04	-0.01	-0.14	-0.04
$\gamma_{12}(1)$	0	0.0	0.0	0.01	0.02	0.02	0.25	0.0	0.0	0.05	-0.01
$\gamma_{12}(2)$	0	0.0	0.0	0.02	0.02	0.02	0.0	0.02	0.0	0.02	0.01
sk_1	0	-0.01	0.01	-0.01	-0.01	-0.01	0.0	0.01	-0.02	-0.01	0.02
sk_2	0	0.01	-0.02	0.01	0.01	-0.01	0.0	0.02	-0.01	0.0	0.05
$kurt_1$	0	-0.04	-0.04	-0.02	-0.03	-0.01	0.0	0.08	-0.09	2.15	7.92
$kurt_2$	0	-0.06	-0.06	-0.04	0.02	-0.03	0.01	0.09	-0.10	1.11	7.68

Table 1b. True cycles' moments and the moments of the extracted cycles

DGP: Distinct stochastic trends

	True	QT	LT	HPF	BPF	Cat	FOD	West	FDF	MED	LWR
μ_1	0	0.0	0.0	0.0	0.0	-0.01	0.50	-0.01	0.0	0.0	0.50
μ_2	0	0.0	0.0	0.0	0.0	-0.01	0.50	0.0	0.0	0.0	0.50
$\gamma_1(0)$	0.94	9.32	14.1	2.24	2.12	2.16	1.77	2.21	0.09	0.88	0.40
$\gamma_1(1)$	0.56	8.35	13.0	1.37	1.26	1.29	-0.28	2.05	0.08	0.19	0.28
$\gamma_1(2)$	0.44	7.68	12.3	0.87	0.75	0.80	-0.03	1.85	0.04	-0.03	0.12
$\gamma_2(0)$	0.94	9.61	13.7	2.19	2.13	2.14	1.77	2.19	0.09	0.89	0.40
$\gamma_2(1)$	0.56	8.64	12.9	1.33	1.26	1.28	-0.28	2.05	0.08	0.20	0.28
$\gamma_2(2)$	0.44	7.96	12.1	0.83	0.76	0.80	-0.03	1.86	0.04	-0.04	0.12
$\gamma_{12}(-2)$	0	0.16	-0.49	0.01	0.02	0.01	0.0	0.01	0.0	0.02	0.01
$\gamma_{12}(-1)$	0	0.17	-0.49	0.01	0.02	0.01	0.24	0.0	0.0	0.04	-0.01
$\gamma_{12}(0)$	-0.25	-0.07	-0.72	-0.23	-0.23	-0.23	-0.49	0.0	-0.01	-0.15	-0.04
$\gamma_{12}(1)$	0	0.17	-0.45	0.03	0.02	0.01	0.24	0.0	0.0	0.04	-0.01
$\gamma_{12}(2)$	0	0.17	-0.42	0.03	0.02	0.0	-0.01	0.01	0.0	0.02	0.01
sk_1	0	0.0	0.01	0.0	0.01	-0.04	0.0	-0.02	-0.02	0.02	0.01
sk_2	0	0.01	0.01	0.0	0.01	-0.01	0.0	-0.02	-0.01	0.02	-0.06
$kurt_1$	0	-0.29	-0.39	-0.07	-0.06	0.09	0.03	1.26	-0.09	1.57	2.47
$kurt_2$	0	-0.30	-0.38	-0.04	-0.03	0.05	-0.01	1.32	-0.11	1.63	2.72

Table 1c. True cycles' moments and the moments of the extracted cycles

DGP: Common stochastic trend correlated with the cycles

	True	QT	LT	HPF	BPF	Cat	FOD	West	FDF	MED	LWR
μ_1	0	0.0	0.0	0.0	0.0	-0.02	0.49	0.01	0.0	0.0	0.49
μ_2	0	0.0	0.0	0.0	0.0	-0.02	0.49	0.02	0.0	0.0	0.49
$\gamma_1(0)$	0.94	9.8	14.6	2.65	2.58	2.59	2.76	2.11	0.10	1.41	0.48
$\gamma_1(1)$	0.56	8.4	13.1	1.29	1.22	1.24	-0.77	1.94	0.08	0.18	0.30
$\gamma_1(2)$	0.44	7.7	12.4	0.79	0.73	0.76	-0.04	1.72	0.04	-0.07	0.11
$\gamma_2(0)$	0.94	8.9	13.6	1.73	1.67	1.72	0.77	2.58	0.08	0.43	0.31
$\gamma_2(1)$	0.56	8.4	13.1	1.37	1.31	1.35	0.21	2.44	0.07	0.22	0.25
$\gamma_2(2)$	0.44	7.7	12.4	0.87	0.81	0.85	-0.03	2.56	0.04	0.01	0.13
$\gamma_{12}(-2)$	0	7.8	14.5	1.10	1.04	1.07	-0.01	1.09	0.0	0.15	0.21
$\gamma_{12}(-1)$	0	8.4	13.1	1.57	1.51	1.54	0.74	1.18	0.03	0.43	0.28
$\gamma_{12}(0)$	-0.25	8.3	13.0	1.37	1.30	1.36	0.49	1.29	0.06	0.32	0.25
$\gamma_{12}(1)$	0	7.5	12.2	0.72	0.66	0.71	-0.26	1.33	0.07	-0.02	0.12
$\gamma_{12}(2)$	0	6.9	11.6	0.40	0.34	0.39	-0.01	1.29	0.06	-0.11	0.02
sk_1	0	0.0	0.01	-0.01	0.01	-0.01	-0.01	-0.03	-0.01	-0.01	0.03
sk_2	0	0.02	0.01	0.0	0.02	-0.04	-0.01	-0.04	-0.02	-0.03	0.01
$kurt_1$	0	-0.24	-0.36	-0.03	-0.02	0.04	-0.03	1.25	-0.10	1.10	4.86
$kurt_2$	0	-0.30	-0.41	-0.05	-0.03	0.11	-0.01	1.22	-0.10	3.30	0.42

Table 2a. Comparative performance of SETs

DGP: Two deterministic trends

(Numbers in brackets stand for standard deviations)

Filter	ADM	Filter	ADME	Filter	ADME2	Filter	Theil
QT	0.052 (0.024)	QT	0.048 (0.021)	QT	0.030 (0.022)	QT	0.110 (0.034)
LT	0.081 (0.027)	LT	0.075 (0.024)	LT	0.054 (0.026)	LT	0.152 (0.034)
HPF	0.190 (0.050)	HPF	0.164 (0.037)	HPF	0.170 (0.047)	HPF	0.276 (0.038)
BPF	0.197 (0.053)	BPF	0.169 (0.039)	BPF	0.178 (0.048)	BPF	0.282 (0.037)
CAT	0.199 (0.051)	CAT	0.170 (0.038)	CAT	0.179 (0.046)	CAT	0.291 (0.037)
MixFdw	0.257 (0.055)	MixFdw	0.210 (0.038)	MixFdw	0.225 (0.047)	MixFdw	0.337 (0.035)
MixFew	0.301 (0.058)	MixFew	0.238 (0.038)	MixFew	0.249 (0.043)	MixFew	0.412 (0.032)
West	0.326 (0.060)	West	0.251 (0.037)	FOD	0.261 (0.034)	MED	0.495 (0.031)
FOD	0.330 (0.053)	FDF	0.254 (0.037)	MED	0.262 (0.047)	FOD	0.539 (0.026)
FDF	0.334 (0.060)	FOD	0.258 (0.033)	West	0.279 (0.045)	West	0.775 (0.026)

Table 2b. Comparative performance of SETs

DGP: Common deterministic trend

(Numbers in brackets stand for standard deviations)

Filter	ADM	Filter	ADME	Filter	ADME2	Filter	Theil
QT	0.053 (0.024)	QT	0.050 (0.022)	QT	0.031 (0.025)	QT	0.115 (0.038)
LT	0.123 (0.039)	LT	0.112 (0.033)	LT	0.093 (0.047)	MultHPF	0.180 (0.036)
MultHPF	0.140 (0.047)	MultHPF	0.127 (0.039)	MultHPF	0.135 (0.049)	LT	0.203 (0.029)
MMFdw	0.181 (0.047)	MMFdw	0.161 (0.037)	HPF	0.172 (0.042)	MMFdw	0.210 (0.035)
HPF	0.190 (0.044)	HPF	0.164 (0.034)	BPF	0.177 (0.048)	MMFew	0.262 (0.040)
BPF	0.196 (0.051)	BPF	0.168 (0.038)	MMFdw	0.178 (0.046)	HPF	0.276 (0.036)
CAT	0.201 (0.054)	CAT	0.172 (0.040)	CAT	0.181 (0.050)	BPF	0.280 (0.036)
MMFew	0.218 (0.056)	MMFew	0.190 (0.042)	MMFew	0.206 (0.051)	CAT	0.289 (0.038)
MixFdw	0.253 (0.048)	MixFdw	0.208 (0.033)	MixFdw	0.223 (0.041)	MixFdw	0.337 (0.036)
MixFew	0.317 (0.060)	MixFew	0.248 (0.037)	MixFew	0.256 (0.046)	MixFew	0.425 (0.032)

Table 2c. Comparative performance of SETs

DGP: Two stochastic trends

(Numbers in brackets stand for standard deviations)

Filter	ADM	Filter	ADME	Filter	ADME2	Filter	Theil
MixFdw	0.246 (0.049)	MixFdw	0.200 (0.032)	MixFdw	0.179 (0.035)	MixFdw	0.525 (0.024)
MixFew	0.252 (0.051)	MixFew	0.206 (0.032)	MED	0.184 (0.044)	MixFew	0.527 (0.024)
FDF	0.323 (0.056)	MED	0.247 (0.038)	MixFew	0.187 (0.034)	MED	0.546 (0.021)
MED	0.381 (0.089)	FDF	0.250 (0.035)	FDF	0.271 (0.041)	CAT	0.551 (0.026)
CAT	0.400 (0.100)	CAT	0.285 (0.055)	West	0.296 (0.127)	BPF	0.552 (0.026)
BPF	0.432 (0.101)	West	0.300 (0.105)	CAT	0.330 (0.078)	HPF	0.556 (0.025)
FOD	0.433 (0.049)	BPF	0.301 (0.055)	FOD	0.345 (0.033)	FOD	0.619 (0.024)
HPF	0.470 (0.120)	HPF	0.316 (0.060)	BPF	0.354 (0.078)	West	0.716 (0.062)
West	0.988 (1.321)	FOD	0.323 (0.030)	HPF	0.380 (0.086)	QT	0.736 (0.041)
QT	4.6 (2.0)	QT	0.679 (0.081)	QT	0.894 (0.119)	LT	0.768 (0.043)

Table 2d. Comparative performance of SETs

DGP: Common stochastic trend not correlated with cycles

(Numbers in brackets stand for standard deviations)

Filter	ADM	Filter	ADME	Filter	ADME2	Filter	Theil
MMFdw	0.226 (0.045)	MMFdw	0.187 (0.031)	MMFdw	0.180 (0.032)	MMFdw	0.440 (0.028)
MMFew	0.231 (0.041)	MMFew	0.190 (0.029)	MMFew	0.181 (0.029)	MMFew	0.444 (0.029)
MixFdw	0.312 (0.061)	MixFdw	0.247 (0.036)	MED	0.228 (0.043)	MixFdw	0.522 (0.023)
MixFew	0.313 (0.058)	MixFew	0.249 (0.034)	MixFdw	0.251 (0.038)	MixFew	0.524 (0.023)
FDF	0.335 (0.062)	FDF	0.257 (0.038)	MixFew	0.254 (0.039)	MultHPF	0.528 (0.036)
MED	0.424 (0.092)	MED	0.276 (0.037)	FDF	0.280 (0.045)	MED	0.542 (0.020)
FOD	0.465 (0.052)	FOD	0.341 (0.032)	FOD	0.374 (0.032)	CAT	0.552 (0.027)
CAT	0.631 (0.168)	West	0.358 (0.189)	West	0.384 (0.248)	BPF	0.552 (0.031)
BPF	0.673 (0.185)	CAT	0.410 (0.073)	CAT	0.525 (0.106)	HPF	0.557 (0.032)
HPF	0.740 (0.199)	BPF	0.426 (0.077)	BPF	0.551 (0.111	FOD	0.617 (0.022)

Table 2e. Comparative performance of SETs

DGP: Common stochastic trend correlated with cycles

(Numbers in brackets stand for standard deviations)

Filter	ADM	Filter	ADME	Filter	ADME2	Filter	Theil
MMFew	0.279 (0.046)	MMFdw	0.219 (0.032)	MMFew	0.225 (0.033)	MMFdw	0.457 (0.030)
MMFdw	0.281 (0.052)	MMFew	0.219 (0.030)	MMFdw	0.227 (0.031)	MMFew	0.457 (0.030)
FDF	0.330 (0.060)	FDF	0.255 (0.038)	FDF	0.278 (0.044)	MultHPF	0.531 (0.032)
MixFdw	0.354 (0.061)	MixFdw	0.269 (0.036)	MED	0.284 (0.036)	MixFdw	0.540 (0.022)
MixFew	0.359 (0.061)	MixFew	0.275 (0.036)	MixFdw	0.284 (0.039)	MixFew	0.546 (0.021)
FOD	0.507 (0.056)	MED	0.322 (0.033)	MixFew	0.292 (0.038)	CAT	0.556 (0.023)
MED	0.564 (0.129)	FOD	0.340 (0.033)	West	0.371 (0.244)	BPF	0.556 (0.027)
CAT	0.641 (0.175)	West	0.351 (0.188)	FOD	0.374 (0.038)	HPF	0.560 (0.030)
BPF	0.666 (0.160)	CAT	0.403 (0.075)	CAT	0.513 (0.112)	MED	0.566 (0.020)
HPF	0.737 (0.199)	BPF	0.415 (0.069)	BPF	0.532 (0.101	FOD	0.645 (0.022)

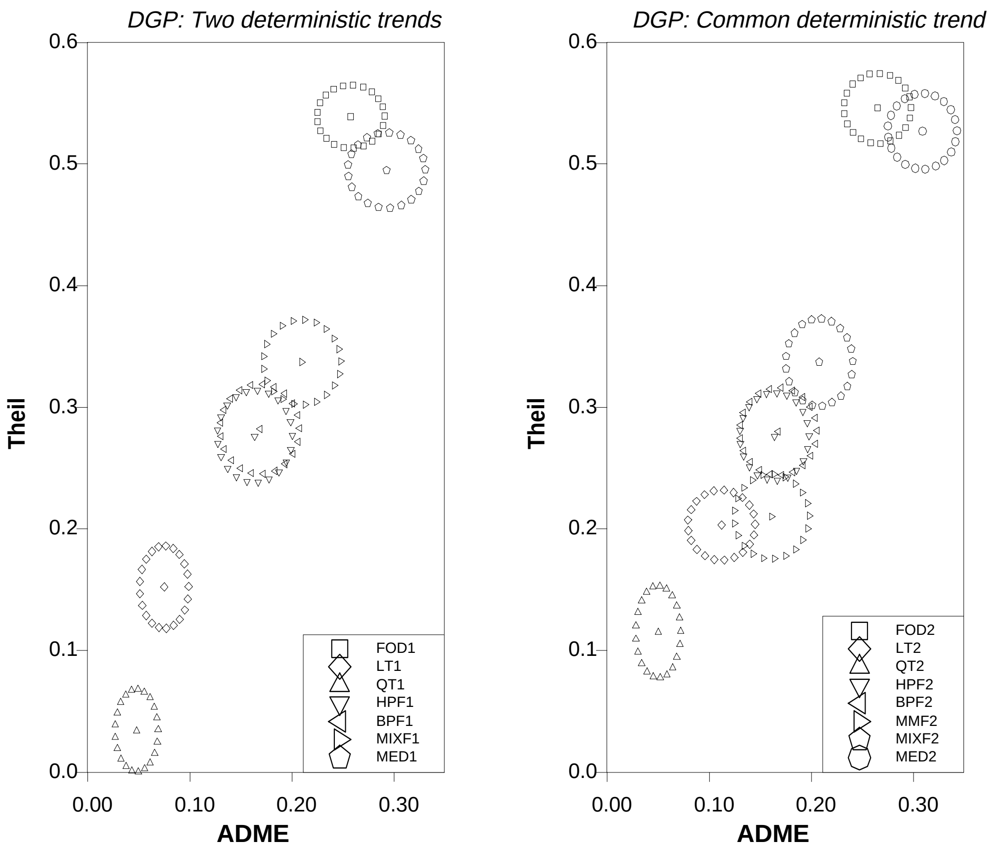

Figure 1: Comparative performance of SETs for the processes with deterministic trends

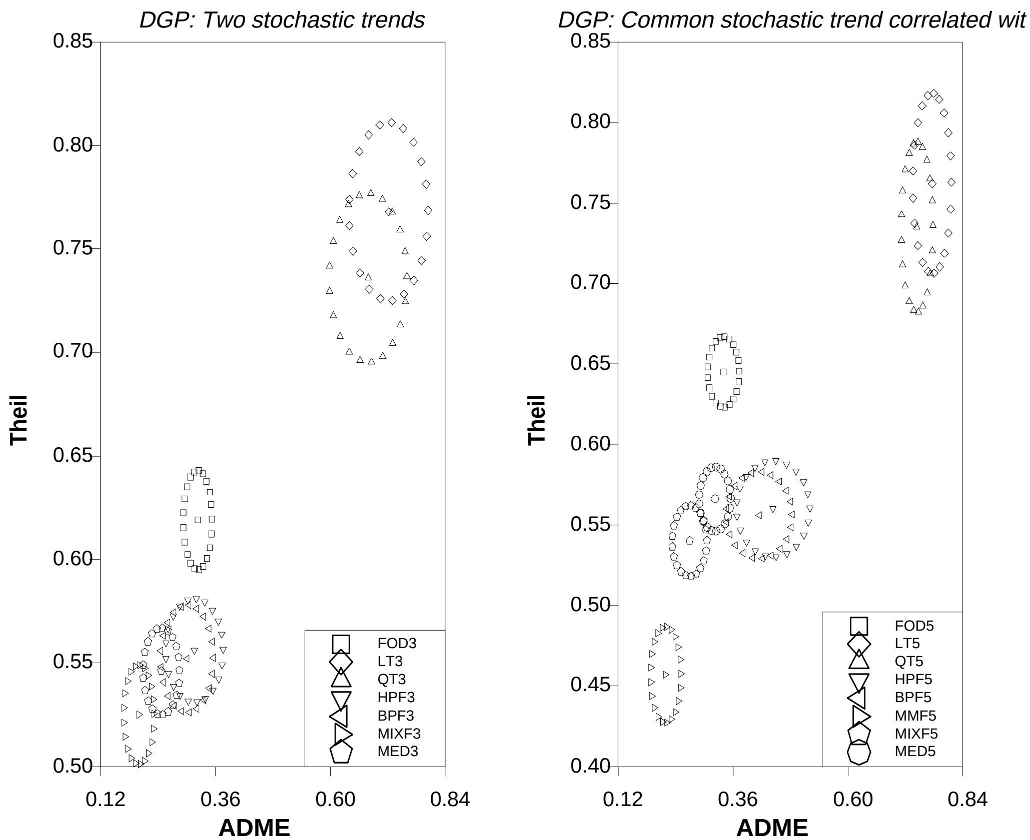

Figure 2: Comparative performance of SETs for the processes with stochastic trends