CEIS Tor Vergata
Centre for International Studies on Economic Growth
TOR VERGATA

CEIS Tor Vergata
RESEARCH PAPER SERIES

Working Paper No. 105 May 2007

Band Spectral Estimation for Signal Extraction

Tommaso Proietti

CEIS Tor Vergata - Research Paper Series, Vol. 35, No. 105, May 2007

Band Spectral Estimation for Signal Extraction

Abstract

The paper evaluates the potential of band spectral estimation for extracting signals in economic time series. Two situations are considered. The first deals with trend extraction when the original data have been permanently altered by routine operations, such as prefiltering, temporal aggregation and disaggregation, and seasonal adjustment, which modify the high frequencies properties of economic time series. The second is when the measurement model is only partially specified, in that it aims at fitting the series in a particular frequency range, e.g. at interpreting the long run behaviour. These issues are illustrated with reference to a simple structural model, namely the random walk plus noise model.

Keywords: Temporal Aggregation, Seasonal Adjustment, Trend Component, Frequency Domain.

JEL classification: C22, E3.

1 Introduction

Many signal extraction problems in economics deal with the measurement of the underlying evolution of economic variables, such as inflation and output. These problems can be addressed in a parametric framework, whereby a measurement model describing the evolution of the signal is specified and estimated according to some criterion; given the parameter estimates, suitable algorithms, such as the Wiener-Kolmogorov filter or the Kalman filter and smoother (KFS), carry out the necessary computations for estimating the signal.

The specification issue will not be considered by this paper, which rather takes the parametric model as fixed. In particular, we shall focus on a simple model breaking down the series into a random walk trend and a white noise component, mutually orthogonal, which is known as the *random walk plus noise* (RWpN) model. This model owes its success to its simplicity, as it provides a basic separation between transitory and permanent dynamics depending on a single smoothness parameter, which determines the weights that are assigned to the available observations for forecasting and trend estimation.

The paper deals with the estimation of the smoothness parameter by band spectral (BS) maximum likelihood (ML). The latter is a variant of ML in the frequency domain which assigns different weights to the Fourier frequencies and it has a fairly long and well established tradition in the time series and econometric literature.

The possibility of eliminating a fixed number of Fourier frequencies is mentioned in Hannan (1969, p. 584), who considers the case when a band is affected by noise. The marginal likelihood (with respect to the drift parameter), which amounts to removing the zero frequency term from the standard likelihood, can also be seen as an extreme particular case (Shephard, 1993, sec. 5) of BS estimation.

The idea of differential weighting according to the frequency is considered in Robinson (1977) and Thomson (1986). The former considers the case when ”omission of small clusters” of frequencies ”at strategic places may reduce seasonal noise” or measurement noise (p. 181). Another motivation is when the underlying process is band limited, that is it has nonzero spectral density only in a specific frequency range (Robinson, p. 180). This case is considered in detail by Pollock (2006a, 2006b), who also provides illustrations with reference to the problem of trend extraction in economic time series. Thomson (1986) adopts BS estimation for the purpose of estimating ARMA models that are noninvertible at a fixed number of frequencies; noninvertibility may result from stationarity inducing transformations, such as differencing.

Moreover, BS estimation has successful applications in regression, when a behavioural relationship is not stable across the frequency range, e.g. the price elasticity with respect to wages is larger in the long run (low frequencies) than in the short run (high frequencies); see Engle (1978) and Jaeger (1992). The theory of BS regression is set forth in Engle (1974), whereas Corbae, Oularis and Phillips (2002) extend it when stochastic and deterministic trends are present.

A similar idea is present in Watson (1993), Diebold, Ohanian and Berkowitz (1998), and Christiano and Vigfusson (2003). These references advocate the use of BS techniques for estimating dynamic structural economic models that are only partially specified, as it occurs when the model explicitly formulates long run behavioural relationships, but is otherwise agnostic on the short run behaviour. They also consider the assessment of goodness of fit in the frequency domain, namely the use of BS methods for evaluating how well a parametric model fit the data spectra across a specified frequency range.

In general, BS estimation is close to the notion of a local likelihood as adopted in cross-sectional (generalised) linear models with independently distributed observations; see, for instance, Hastie, Tibshirani and Friedman (2001, sec. 6.5), and the references therein.

The objective of this paper is to evaluate the role of BS estimation for extracting the trend from an economic time series; our applications will deal with the measurement of core inflation and potential output. There are two situations in which BS proves valuable.

The first is trend extraction when the original data have been permanently altered by routine operations, such as prefiltering, temporal aggregation and disaggregation, and seasonal adjustment, which modify the high frequencies properties of economic time series. In this situation the RWpN is correctly specified, but the original data have been transformed to an extent that they ”lack power” at the high frequencies. BS estimation will aim at discounting those Fourier frequencies which are most affected by preliminary transformations.

The second arises when the RWpN model is used as an approximation to an unknown true model separating the long run and the short run. For instance, we may be interested in extracting a random walk trend from an economic time series, leaving the short run unspecified. The approximate model can only provide a stripped to the bone representation of economic dynamics, so that for real output series such as gross domestic product (GDP) at constant prices we expect that the transitory component has richer dynamics than pure white noise, as postulated by the RWpN model.

Hence, the second situation differs from the first in that we assume that the RWpN model is misspecified or only partially specified. In the time series literature considerable attention has been attracted by the issue of determining estimation criteria alternative to ML that could guarantee efficiency in prediction and signal extraction from a model that would otherwise be rejected from the standard model building methodology.

We aim at illustrating that BS estimation provides a valuable tool for signal extraction from a misspecified model: using the periodogram ordinates in a neighbourhood of the zero frequency, thereby emphasising the separation of the long-run behaviour from the short run, it allows the extraction of trends from an economic time series using a simple and well understood model.

In this respect, its scope is similar to multistep (or adaptive) estimation of the RWpN, which is discussed in Cox (1961), Tiao and Xu (1993) and Haywood and Tunicliffe Wilson (1997), thereby providing a valid alternative estimation tool.

The paper is structured as follows: in the next section we set off by introducing and reviewing the properties of the our workhorse model, the RWpN model. Estimation in the frequency domain and band spectral estimation are the topic of section 3. Section 4 introduces the scope of BS estimation using an artificial example. Temporal disaggregation and seasonal adjustment are deal with in sections 5 and 6, respectively. The use of BS for signal extraction under model misspecification is illustrated in section 7. Section 8 concludes the paper.

2 The random walk plus noise model

The RWpN model provides the following decomposition of a univariate time series, y_t, into trend and irregular components, denoted respectively by μ_t and ϵ_t:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, \qquad t = 0, 1, \ldots, T, \qquad \epsilon_t \sim \text{NID}(0, \sigma^2_\epsilon), \\ \mu_t &= \mu_{t-1} + \beta + \eta_t, \qquad\qquad\qquad \eta_t \sim \text{NID}(0, \sigma^2_\eta). \end{aligned} \tag{1}$$

The trend is a random walk with normal and independently distributed (NID) increments, whereas the irregular is a Gaussian white noise (WN) process. When the drift is absent, i.e. $\beta = 0$, the model is also known as the *local level model*, see Harvey (1989). We assume throughout that $\text{E}(\eta_t \epsilon_{t-j}) = 0$ for all t and j, so that the components are orthogonal.

The RWpN has a long tradition and a well established role in the analysis of economic time series, as it provides the model based interpretation for the celebrated forecasting technique known as *exponential smoothing*, which is widely used in applied economic forecasting and fares remarkably well in forecast competitions; see Muth (1960) and the comprehensive review by Gardner (1985). For later treatment it is helpful to review the main properties of the model. Further details can be found in Harvey (1989) and Durbin and Koopman (2001).

The reduced form representation of (1) is an IMA(1,1) model: $\Delta y_t = \beta + \xi_t - \theta \xi_{t-1}, \xi_t \sim \text{NID}(0, \sigma^2)$, with the MA parameter subject to the restriction $0 \leq \theta \leq 1$. Equating the autocovariance generating functions of Δy_t implied by the IMA(1,1) and the structural representation (1), it is possible to establish that $\sigma^2_\eta = (1-\theta)^2\sigma^2$ and $\sigma^2_\epsilon = \theta\sigma^2$. The signal-noise ratio (SNR), $q = \sigma^2_\eta/\sigma^2_\epsilon$, depends uniquely on θ: $q = (1-\theta)^2/\theta$. The reciprocal of the SNR is a measure of relative smoothness of the trend: if q^{-1} is large, then the trend varies little with respect to the irregular, and thus it can be regarded as ”smooth”.

The rescaled parametric spectrum of Δy_t, denoted $g(\omega)$, where ω is the frequency in radians, defined in the interval $\in [0, \pi]$, is given by

$$g(\omega) = \sigma^2_\eta + 2(1 - \cos\omega)\sigma^2_\epsilon = (1 + \theta^2 - 2\theta\cos\omega)\sigma^2,$$

and has a maximum at the π frequency.

Assuming a doubly infinite sample, the one-step-ahead predictions, the filtered and smoothed estimates of the trend component are given, respectively, by:

$$\tilde{\mu}_{t+1|t} = \tilde{\mu}_{t|t} = (1-\theta)\sum_{j=0}^{\infty}\theta^j y_{t-j}, \quad \tilde{\mu}_{t|\infty} = \frac{1-\theta}{1+\theta}\sum_{j=-\infty}^{\infty}\theta^{|j|}y_{t-j};$$

these expressions follow from applying the Wiener-Kolmogorov prediction and signal extraction formulae, see Whittle (1983) and Harvey and Koopman (2000). In finite samples the computations are performed by the KFS.

The parameter θ (or, equivalently, the SNR q) is essential in determining the weights that are attached to the observations for signal extraction and prediction. When $\theta = 0$, y_t is a pure random walk and thus the estimates of the signal are as local as possible, $\tilde{\mu}_{t+1|t} = \tilde{\mu}_{t|t} = \tilde{\mu}_{t|\infty} = y_t$; on the contrary, when $\theta = 1$, the trend is as smooth as possible, being a straight line passing through the observations.

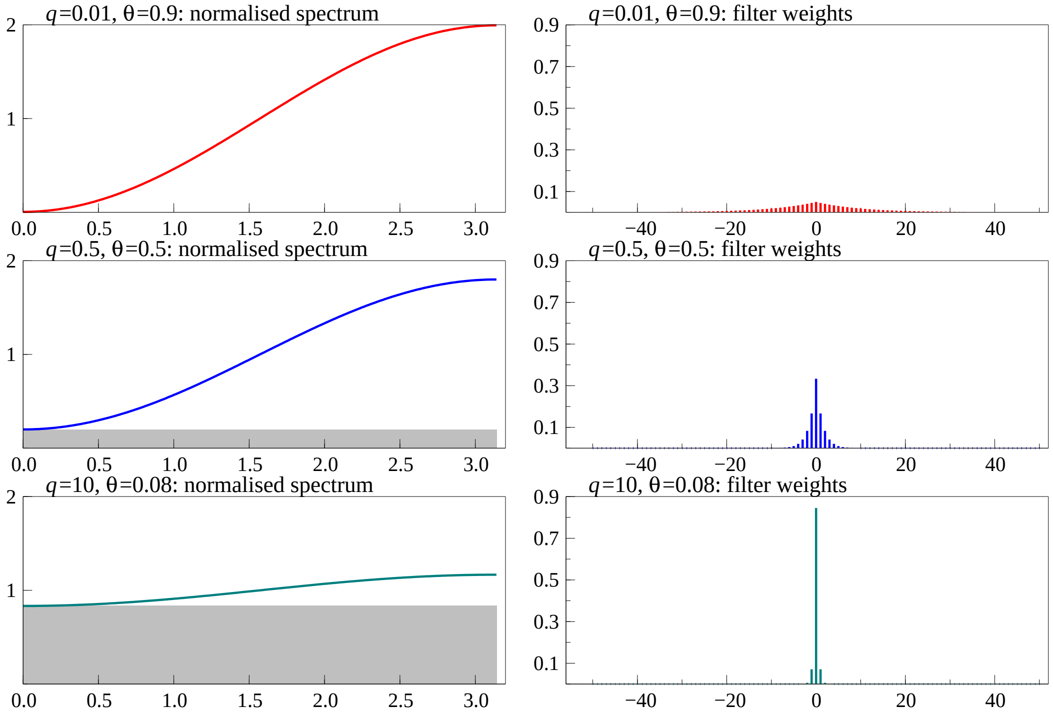
Figure 1 displays in the right panels the normalised spectral density of Δy_t, $g(\omega)/(1+\theta^2)$, for three values of the parameter q. The area beneath the curve is equal to one. The height of the rectangular box is the ratio of σ^2_η to the variance of Δy_t and represents the portion of the total variability is due to trend movements. As q increases (θ decreases) the spectrum tends to a horizontal straight line. The two-sided filter weights for trend estimation for a sequence of 101 observations corresponding to the different values of the SNR parameter are depicted in the right panels. The weights decline symmetrically and geometrically at different rates, depending on the value of the SNR (or θ). As q increases the trend estimates become more and more local until the weighting function is unity for the current observation and zero elsewhere, which occurs when y_t is a random walk.

3 Band spectral estimation

Estimation of the unknown parameters, $\theta \in [0,1]$ and σ^2 (or equivalently q and σ^2_ϵ) can be done in the frequency domain. Maximum likelihood estimation is then based on the stationary representation of the model, $\Delta y_t = \beta + \eta_t + \Delta\epsilon_t$, or in terms of the reduced form, $\Delta y_t = \beta + \xi_t - \theta\xi_{t-1}$, $\theta \in [0,1]$.

While the time domain likelihood is evaluated via a recursive orthogonalisation, known as the prediction error decomposition, performed by the Kalman filter, the frequency domain likelihood

Figure 1: Normalised spectral density of Δy_t (left panels) and two-sided filter weights attributed to $y_{t-j}, j = 0, \pm 1, \ldots$, for signal extraction by the RWpN model with different SNR. The horizontal axis measures the lag j (right panels).



is based on an alternative orthogonalisation, achieved through a Fourier transform. We refer to Nerlove, Grether and Carvalho (1995) and Harvey (1989, sec. 4.3) for a comprehensive treatment on the derivation of the likelihood function and the nature of the approximation involved.

Given the availability of the differenced observations $\Delta y_t, t = 1, 2, \ldots, T$, let us denote the Fourier frequencies by $\omega_j = \frac{2\pi j}{T}$, $j = 0, 1, \ldots, (T-1)$. Apart from a constant, the Whittle's approximation to the likelihood function is defined as follows:

$$\text{loglik} = -\frac{1}{2}\sum_{j=0}^{T-1}\left[\log g(\omega_j) + 2\pi\frac{I(\omega_j)}{g(\omega_j)}\right] \tag{2}$$

where $g(\omega_j)$ is the scaled parametric spectrum of the stationary representation of the RWpN model evaluated at frequency ω_j, that is $g(\omega_j) = \sigma^2(1 + \theta^2 - 2\theta\cos\omega_j)$, and $I(\omega_j)$ is the sample spectrum:

$$I(\omega_j) = \frac{1}{2\pi}\left[c_0 + 2\sum_{\tau=1}^{T-1} c_\tau \cos(\omega_j \tau)\right],$$

where c_τ is the sample autocovariance of Δy_t at lag τ. Notice that for $1 \leq j \leq [T/2]$, where $[T/2] = T/2$, if T is even, and $[T/2] = (T-1)/2$ otherwise, $I(\omega_j) = I(\omega_{T-j})$ and $g(\omega_j) = g(\omega_{T-j})$, that is both the sample and the parametric spectrum are symmetric around π. Hence, in the sequel these quantities will be plotted only in the range $[0, \pi]$.

Band spectral estimation is a particular form of weighted likelihood estimation in the frequency domain, where the objective function (2) is modified as follows:

$$\text{loglik}_{(BS)} = -\frac{1}{2}\sum_{j=0}^{T-1} \mathrm{w}_j \left[\log g(\omega_j) + 2\pi\frac{I(\omega_j)}{g(\omega_j)}\right]. \tag{3}$$

In the above expression w_j is a nonnegative weight, depending on the Fourier frequency, which is symmetric with respect to π, implying $\mathrm{w}_j = \mathrm{w}_{T-j}$.

Typically, w_j takes the value of 1 in a certain band and 0 elsewhere, so that (3) is the summation over the frequencies of interest. The simplest option is to consider the uniform weighting function, defined as follows:

$$\mathrm{w}_j = \mathrm{w}_{T-j}, \quad j = 1, \ldots, [T/2], \tag{4}$$

$$\mathrm{w}_j = \begin{cases} 1, & \omega_j \leq \omega_c \quad j < [T/2] \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

The frequency ω_c is the cutoff frequency in the range $[0, \pi]$. Equation (4) establishes that the weighting pattern is symmetric around the π frequency.

As hinted in the Introduction, there are a variety of motivations for using (3) in place of the standard likelihood (2). We mention in passing that we do not explore further a special case for which BS estimation seem to be well suited, which arises when the trend disturbance is a band limited white noise process, characterised by the following parametric spectrum (up to a factor of proportionality):

$$g(\omega) = \begin{cases} \sigma_\eta^2 + 2(1-\cos\omega)\sigma_\epsilon^2, & \omega \leq \omega_c \\ 2(1-\cos\omega)\sigma_\epsilon^2, & \omega \leq \omega_c \end{cases}$$

In this particular case, we could fit the RWpN model using (3) with (4)-(5) and cutoff ω_c.

A first motivation for downweighting the high frequencies, that will be illustrated in the next three sections, arises when certain transformation of the series have taken place beforehand, so that the original amplitude of the frequency components in the series has been modified to an extent that the condition under which the trend-irregular decomposition of y_t is admissible is no longer met. In terms of the MA(1) reduced form representation for Δy_t, the condition requires that θ is in $[0, 1]$, which implies that the spectral density of Δy_t is a minimum at the zero frequency (see figure 1).

Moreover, if our interest lies in long range forecasting and in the estimation of long-run trends, it is of reduced importance how well the model fits the sample spectrum at the high frequencies, which express uninteresting fluctuations from the analyst perspective, such as trading day variation and other short lived components.

4 An artificial example

In this section we provide an artificial example that motivates the adoption of BS estimation. The first panel of Figure 2 displays the monthly inflation rate for the Consumer Price Index (CPI) for France (all items, base year 2000), for the sample period going from February 1970 to December 2004.

If the RWpN model (with no drift) is estimated according to (2), the ML estimate of θ amounts to 0.7 (the estimated SNR is 0.13). The plot also shows the estimated trend in inflation, e.g. a univariate *core inflation* measure, using the Kalman filter and smoother (see Durbin and Koopman, 2001, for details), which is much smoother than the original series. All the computations were carried out using Ox 3.20 by Doornik (2001) and the package SsfPack 2.2 (Koopman, Shephard and Doornik, 1999) for KFS. The sample spectrum $2\pi I(\omega_j)$ and the fitted parametric spectrum are plotted in the bottom left panel.

Assume now that the series has been prefiltered by the unweighted moving average $(1 + L + L^{-1})/3$, so the series y_t is replaced by the average across three months: $y_t^* = (y_{t-1}+y_t+y_{t+1})/3$. A three months moving average is often used by statistical agencies for publishing time series characterised by short-term volatility. The transformed series and its sample spectrum are plotted in the right panels of Figure 2. It can be seen that the sample spectrum has now much reduced power at the higher frequencies; this is so since the lower frequencies are preserved to a higher extent by the moving average filter.

If the RWpN is fitted to the transformed series, the maximiser of (2) is $\hat{\theta} = 0$, which implies $\hat{\sigma}_\epsilon^2 = 0$, and $\tilde{\mu}_{t|T} = y_t^*$. As a result the estimated trend is coincident with the series y_t^*; in the light of the previous evidence, this is too rough and volatile as a measure of trend. The sample spectrum of Δy_t^* is interpolated by a constant spectrum, as it is evident from the bottom right plot, which also shows that the random walk model fits the low frequencies rather poorly.

In fact, the transformation by a moving average has altered the high frequency dynamics in the series, while the long run properties are retained to a great extent. A plot of the gain of the filter would make this clear. As a result the signal estimates are more erratic than they should be and all the movements in the series are considered as permanent.

We consider now the BS estimation of the RWpN model using (3) with (4)-(5) as a weighting function and different cutoffs frequencies: $\omega_c = \pi/k, k = 1, \dots, 18$. Obviously, when $\omega_c = \pi$ we recover the standard MLE which maximises (2).

BS estimation is carried out for both the original and the transformed series. Figure 3 displays the BS estimates of θ against the cutoff frequency. As far as the original series is concerned, the latter are very stable and close to the standard MLE ($\tilde{\theta} = 0.7$) for cutoff frequencies corresponding to values of k up to 7, and then they suddenly rise up to 1, in which case the trend extracted by the RWpN model is constant.

Figure 2: Monthly CPI inflation rate, France 1970.2-2004.12. Standard ML estimation of the RWpN model in the frequency domain using the original series (left panels) and the transformed series by a 3-terms centred arithmetic average.

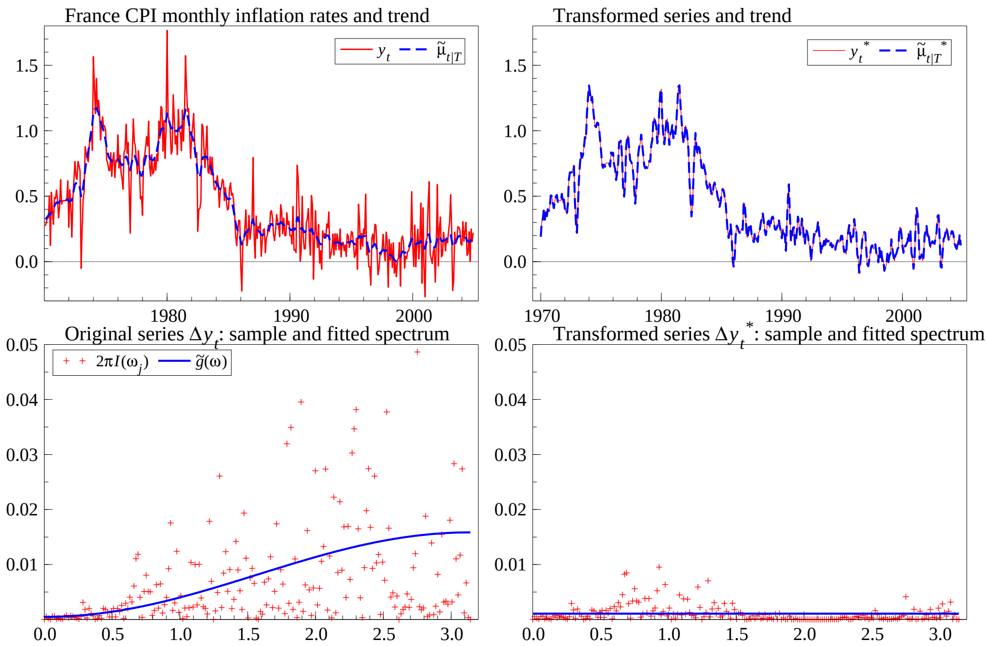
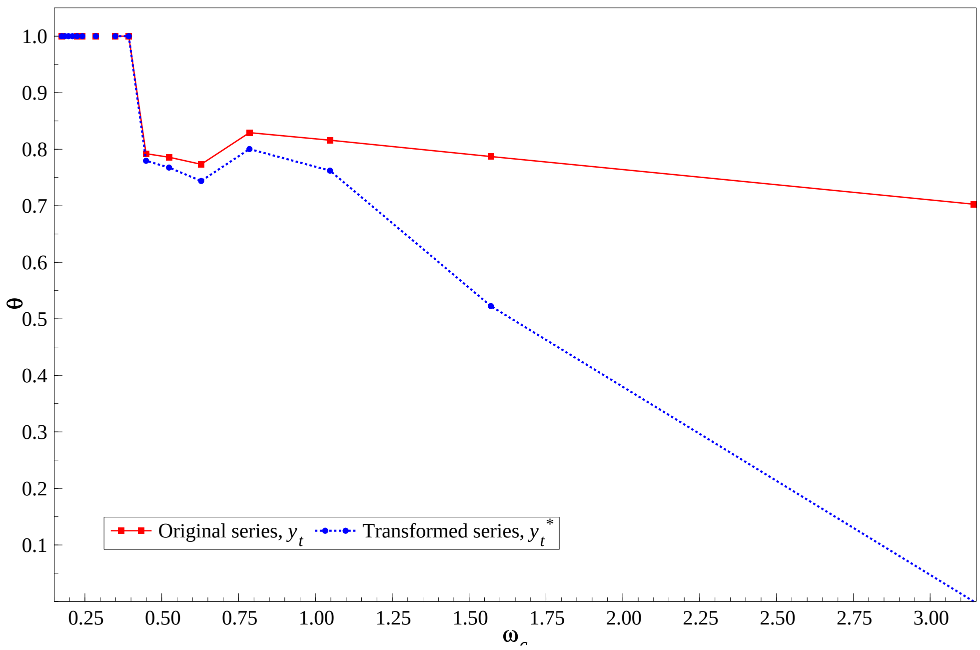


Figure 3: Band spectral maximum likelihood estimates of the parameter θ versus the cutoff frequency, ω_c.



When we apply BS estimation to the transformed series the noticeable fact is that if the cutoff frequency is $\pi/2$, which amounts to ignoring in the estimation all the frequencies corresponding to fluctuations with periodicity smaller than four observations, the MLE of θ moves away from zero, rising up to 0.5. If we further lower the cutoff, we get very close to the BS estimates obtained for the original time series and to the standard MLE for the same series.

This trivial example shows that BS can improve the estimation of trends for series that have been altered by filtering operations. We note in passing that preliminary smoothing of certain components of economic aggregates that are particularly noisy (e.g. certain CPI items such as fruit and vegetables) is not uncommon in the data production process adopted by statistical agencies. The instability of the BS estimates using different cutoffs can be considered as a diagnostic device.

5 Temporal disaggregation

Temporal disaggregation methods play an important role for the construction of short term economic indicators. This is certainly the case for a number of European countries (Eurostat, 1999), including France, Italy and Spain, whose national statistical institutes make extensive use of disaggregation methods, among which the Chow-Lin procedure stands out prominently (Chow and Lin, 1971), for constructing the quarterly national economic accounts using the annual figures and a set of indicators available at the quarterly frequency. These considerations apply to the Italian GDP series, that will be dealt with in section 7.

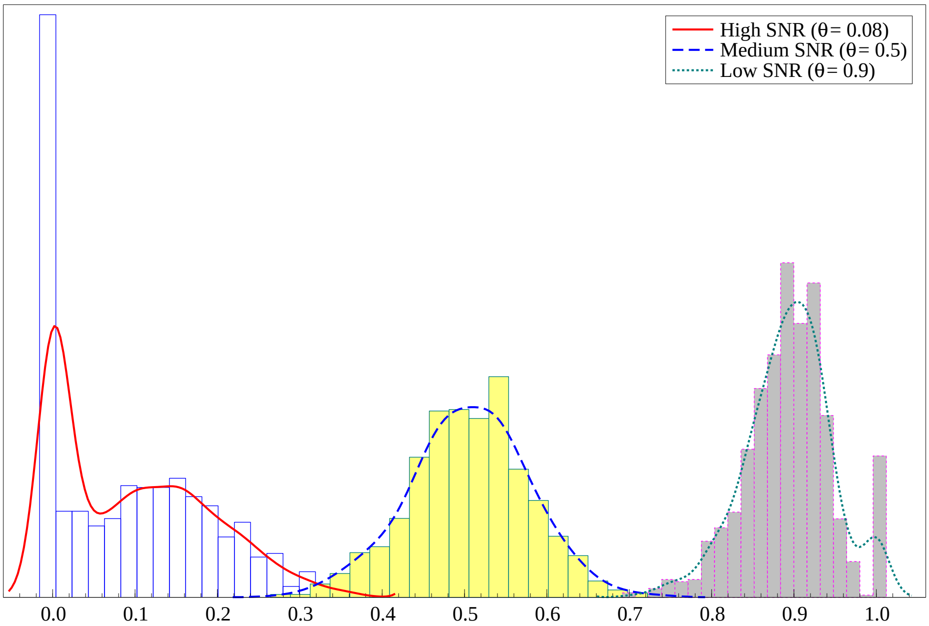
We argue that economic time series resulting from temporal disaggregation suffer from excess smoothness or ”lack of power” at the high frequency. Intuitively, the process of disaggregating aggregate data cannot restore all the power that characterised the original series at the high frequencies. As a consequence, the estimates of the SNR, or the θ parameters will be biased downward, whereas, on the contrary, the contribution of the irregular component will be underestimated.

A Monte Carlo experiment was conducted with the intent of assessing the properties of the standard MLE and BS MLE when the series arises from the disaggregation of a temporal aggregate at a coarser interval, and evaluating the detrimental effects on trend estimation.

The design of the experiment is the following: $M = 1000$ series of length $T = 280$ (e.g. 20 years of monthly observations) are simulated from the true monthly RWpN model (1), using three different values of the SNR: *Low*, $q = 1/100$, corresponding to $\theta = 0.9$; *Medium*, $q = 1/2$, corresponding to $\theta = 0.5$; *High*, $q = 10$, corresponding to $\theta = 0.08$.

The parameter q (or θ) is estimated on the generated series by standard ML in the frequency domain. The distribution of the θ estimates over the 1000 replications is plotted in Figure 4. When the SNR is low, the MLE suffers from what is referred to as the *pile-up problem*, meaning that there is a concentration of density at $\hat{q} = 0$ ($\hat{\theta} = 1$), so that when the true signal has a weak evolution it is estimated as constant. In such cases Shephard (1993) shows that the marginal likelihood (with respect to the initial value of the random walk) alleviates the problem. The marginal likelihood is actually a particular extreme case of (3) since it amounts to setting to zero the weight attached to the zero frequency. When the SNR is high, the *pile-up problem* concerns the value $\theta = 0$, instead, at which the noise contribution is estimated to be zero.

Figure 4: Distribution of MLE of θ for simulated series



Let us assume now that the original series are not available to the analyst, but that the series Y_τ, resulting from temporal aggregation, is available at times $\tau = 1, 2, \dots [n/s]$, where $[n/s]$ denotes the integral part of n/s and Y_τ is generated as follows:

$$Y_\tau = \sum_{j=0}^{s-1} y_{\tau s - j}, \tau = 1, 2, \dots [n/s]. \tag{6}$$

The values of the aggregation interval that we considered were $s = 12$ (e.g. from monthly to annual data; the aggregated series has 20 observations) and $s = 3$ (from monthly to quarterly observations, the aggregated series has 80 observations), but for brevity we report only on the last case, which is fully representative.

Given the availability of the aggregate series alone, let us suppose that Y_τ is disaggregated into a monthly time series in an optimal way, that is using a methodology that returns the minimum mean square estimates of the disaggregate series, denoted y_t^*, satisfying the aggregation constraint $Y_\tau = \sum_{j=0}^{s-1} y_{\tau s-j}^*$.

This is achieved in practice using the state space methods set up in Harvey and Chun (2000) and Proietti (2006). In order to handle temporal aggregation, a new state space representation is derived from that of the underlying true model, an IMA(1,1) model, by augmenting the state vector of the original state space representation with a cumulator variable that is only partially observed. The cumulator is defined as follows:

$$y_t^c = \psi_t y_{t-1}^c + y_t, \quad \psi_t = \begin{cases} 0, & t = s(\tau-1)+1, \tau = 1, \dots, [n/s] \\ 1, & \text{otherwise} \end{cases}$$

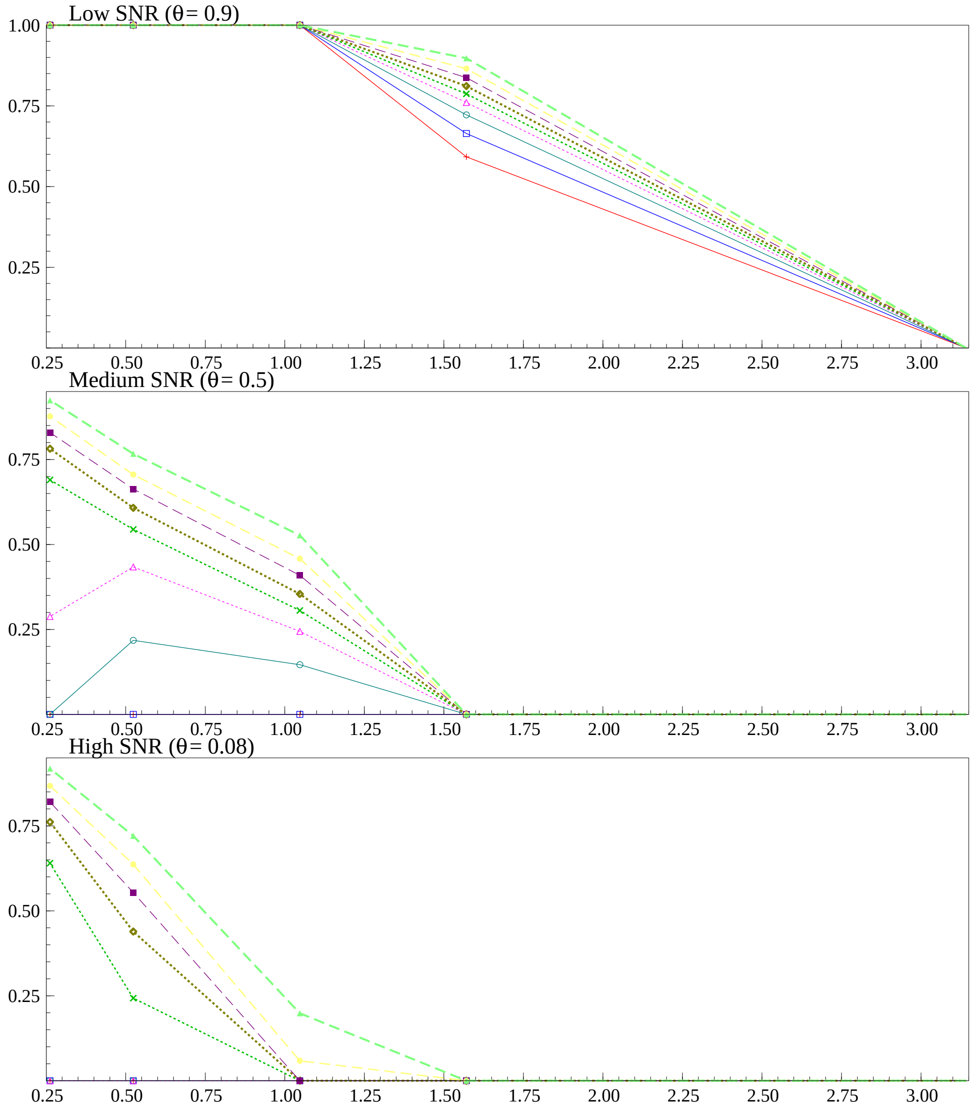
Temporal aggregation is such that only a systematic sample of every s-th value of y_t^c process is observed, $Y_\tau = y_{\tau s}^c, \tau = 1, \dots [n/s]$, whereas all the remaining values are missing. This approach, proposed by Harvey (1989, sec. 6.3), converts the disaggregation problem into a problem of missing values, that can be addressed by skipping certain updating operations in the filtering and smoothing equations. The minimum mean square estimates of the disaggregate series y_t^* are computed by the Kalman filter and smoother applied to the relevant state space model. In setting up the latter, the disaggregation adopts the true parameter values used in the generation of the disaggregate series, and thus there is no parameter uncertainty affecting the estimates y_t^*.

The RWpN model is fitted to the estimated monthly series both by ordinary and band-limited maximum likelihood. The results are summarised in Figure 6, which plots, separately for the three values of the SNR and θ values, the deciles of the distribution of the estimated parameter θ.

The deciles of the standard MLEs correspond to the value of the cutoff frequency $\omega_c = \pi$, which is the outmost value of the frequency range considered. Remarkably, the deciles are zero in the three cases. As a result, regardless of the true SNR, the disaggregated series suffers from excess smoothness and the variance of the irregular is estimated equal to zero. This is inconvenient from the signal extraction perspective: the estimated disaggregated series is smoother than the original simulated series, but yet it is too rough to be comparable to the trend component estimated on the original series.

When we discard the high frequency by using the cutoff $\omega_c = \pi/2$, the distribution of the BS MLE of θ moves away from zero when the true SNR is low ($\theta = 0.9$), as it can be seen from the upper panel of Figure 6. The estimated trend component is much closer to that extracted using the original series. For Medium and High SNR we need lower cutoffs, e.g. $\pi/4$ for the former case, to move the BS estimates closer to the true value and to improve the estimation of trends.

Figure 5: Deciles of the band spectral estimates of the parameter θ for difference cutoff frequencies



5.1 Estimation under temporal aggregation

The previous section illustrated that temporal disaggregation leads to a systematic overestimation of the SNR or, which is the same, to a value of θ that is biased towards 0; the problem is alleviated by BS estimation.

Rather than disaggregating the series and estimating the RWpN model in the estimated series, an alternative strategy is to estimate the RWpN model, which is formulated at the higher observation frequency (e.g. monthly), using directly the aggregate (e.g. quarterly) data.

Assuming as before that aggregation occurs as in (6) and denoting the sample spectrum of $Y_\tau - Y_{\tau-1}, \tau = 1, \dots N$, by $I^\dagger(\omega_j)$, the frequency domain likelihood is

$$\text{loglik} = -\frac{1}{2}\sum_{j=0}^{N-1}\left[\log g_{TA}(\omega_j) + 2\pi\frac{I^\dagger(\omega_j)}{g_{TA}(\omega_j)}\right] \tag{7}$$

where $2\pi g_{TA}(\omega)$ is the parametric spectrum of the aggregate RWpN process. The latter is related to the spectrum of the disaggregate process Δy_t, $2\pi g(\omega)$, by the following expression, which result from the application of the well known *folding* formula to the process $\Delta_s S(L) y_t = S(L)^2 \Delta y_t$, where $S(L) = 1 + L + \cdots + L^{s-1}$ (see Harvey, 1989, sec. 6.3.5):

$$g_{TA}(\omega) = \frac{1}{s}\sum_{h=0}^{s-1}\left|S\left(e^{-\imath\omega_h}\right)\right|^4 g(\omega_h) \tag{8}$$

where $\omega_h = (\omega + 2\pi h)/s$ and

$$\left|S\left(e^{-\imath\omega_h}\right)\right|^2 = S\left(e^{-\imath\omega_h}\right)S\left(e^{\imath\omega_h}\right) = \begin{cases} \frac{1-\cos\omega}{1-\cos\omega_h}, & \omega_h \neq 0 \\ s^2, & \omega_h = 0 \end{cases}$$

Now, there is no way of rewriting (7) as a weighted likelihood. However, minimising

$$\sum_{j=0}^{N-1}\left[2\pi\frac{I^\dagger(\omega_j)}{g_{TA}(\omega_j)}\right] = \sum_{j=0}^{N-1}\mathrm{w}(\omega_j)\left[2\pi\frac{I^\dagger(\omega_j)}{g(\omega_j)}\right]$$

is equivalent to minimising a weighted function, where the weighting function $\mathrm{w}(\omega) = g(\omega)/g_{TA}(\omega)$, is a nondecreasing function of ω, giving more weight to the high frequencies. The latter are weighted more than the other frequencies, depending on the θ value, since they have been subject to compression through aggregation to a larger extent.

In conclusion, estimation of the RWpN model taking into account of aggregation constraints automatically ”redresses the balance” in favour of those frequencies which were affected by aggregation, i.e. the high frequencies. As pointed out also at the end of the next section, BS has a slightly different logic, as it does not attempt to ”redress the balance”, but it selects purposively those frequencies that were not altered by the transformation, and that are likely to provide genuine information on the underlying model.

6 Seasonal adjustment

Signal extraction is often carried out on seasonally adjusted series. This simplifies the formulation and the estimation of the model, since the analyst does not have to specify the seasonal component and estimate the parameters that govern its evolution. Furthermore, it is often the case that seasonally adjusted data are the only available data.

However, seasonal adjustment (SA) can seriously affect the trends estimated from economic time series. As we shall argue in this section, SA provides a further explanation for the ”excess smoothness” of economic time series. We address this issue using a simple Monte Carlo simulation, which is taken from a larger experiment, the full details of which are skipped for brevity.

We generate $M = 1000$ series from the RWpN model augmented by a quarterly seasonal component, $y_t = \mu_t + \gamma_t + \epsilon_t$ where the trend and the irregular are specified as in (1), whereas the seasonal component has the following representation:

$$\gamma_t = \frac{1}{1+L^2}\kappa_{1t} + \frac{1}{1+L}\kappa_{2t},$$

or, equivalently, $(1+L+L^2+L^3)\gamma_t = (1+L)\kappa_{1t} + (1+L^2)\kappa_{2t}$ where $\kappa_{1t} \sim \text{NID}(0, \sigma^2_\omega)$, $\kappa_{2t} \sim \text{NID}(0, \sigma^2_\omega)$, and all the disturbances are mutually uncorrelated.

The parameters are set equal to the following values: $q = 1/2$, $\sigma^2_\omega = 0.1\sigma^2_\epsilon$, $\sigma^2_\epsilon = 1$, $\beta = 0.5$; the true seasonally adjusted series, $y^*_t = y_t - \gamma_t$, has thus a RWpN representation with SNR equal to $1/2$, corresponding to the MA parameter $\theta = 0.5$.

For each of the simulated series we estimate the RWpN model with seasonality by standard MLE in the frequency domain; next, we remove the seasonal component estimated by the KFS, conditional on the estimated parameter values; finally, we estimate the RWpN model (1) on the seasonally adjusted series and compare the parameters and trend estimates with those estimated on the original unadjusted series. We also consider BS estimation, where we delete a range of frequencies around the seasonal frequencies. In particular, $\mathrm{w}_j = 0$ in a neigbourhood $\pm 0.05k$ around the seasonal frequencies $\pi/2, \pi$ and $3\pi/2$, where $k = 1, \ldots, 10$.

Figure 6 illustrates one replication of the Monte Carlo experiment: the simulated series is plotted in the upper panel, along with the true trend and that estimated from the seasonal model. The central panel presents the sample spectrum of the seasonally adjusted series, $\tilde{y}_t^* = y_t - \tilde{\gamma}_t$, where $\tilde{\gamma}_t$ is the minimum mean square estimate of the seasonal component, conditional on the MLE of the parameters $q, \sigma_\epsilon^2, \sigma_\omega^2$ and β, the estimated spectrum of the RWpN model obtained by BS maximum likelihood. The plot also provides the barplot of the 0-1 weights used in (3).

Assuming a doubly infinite sample, the squared gain of the seasonal adjustment filter is:

$$\frac{|S(\omega)|^2\sigma_\eta^2}{|S(\omega)|^2\sigma_\eta^2 + 2(1-\cos\omega)g_\kappa(\omega)} = \frac{1}{1 + \frac{2(1-\cos\omega)}{|S(\omega)|^2}g_\kappa(\omega)} \tag{9}$$

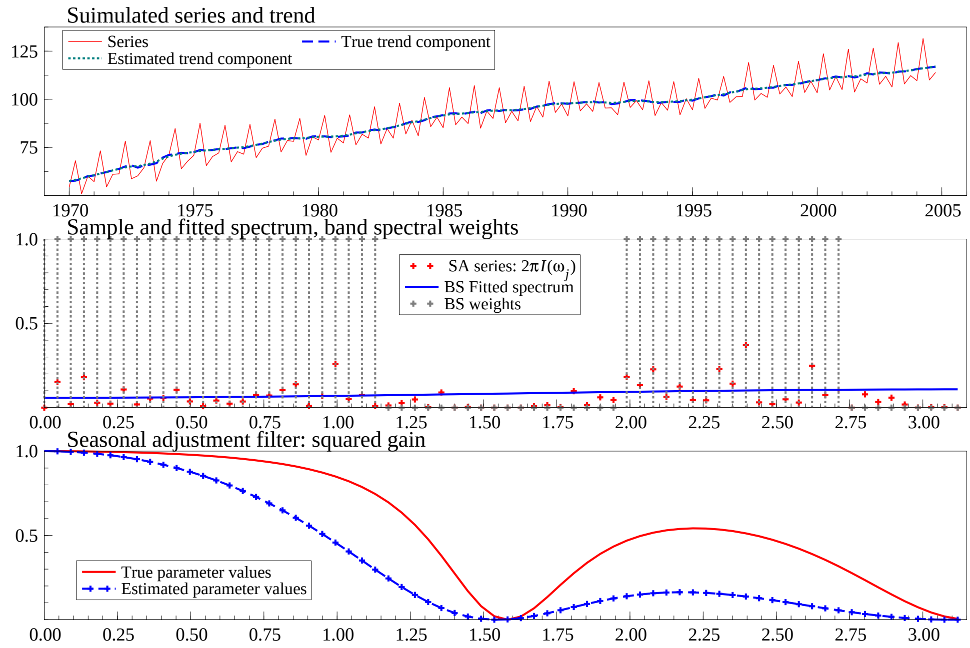
where $g_\kappa(\omega) = 2\left[(1+\cos\omega) + 0.5(1+\cos 2\omega)\right]\sigma_\kappa^2$.

The solid line in the bottom panel of Figure 6 is the squared gain of the theoretical filter, i.e. the above expression evaluated at the true parameter values. The dashed line is the squared gain evaluated at the MLE of the parameters. The gain is zero at the seasonal frequencies and close to zero around them, but the true one dominates the estimated one. As a matter of fact, given the parameter setting, overadjustment is the dominant feature of the adjustment, in that too much power is removed around the seasonal frequencies.

Coming to the properties of the parameters of the RWpN model on the seasonally adjusted series, the evidence can be summarised as follows: standard MLE suffers from a marked tendency to overestimated the SNR; equivalently, θ is biased towards 0. From the signal extraction perspective the estimated trends tend to be coincident with the seasonally adjusted series and are too variable with respect to the original trends. BS estimation improves this situation since it suppresses the frequency that are likely to have suffered most from the adjustment. However, the problem is alleviated, but not resolved and also BS suffers from systematic underestimation of the noise variance.

This is illustrated by Figure 7, which is the scatterplot of the standard MLEs of θ computed

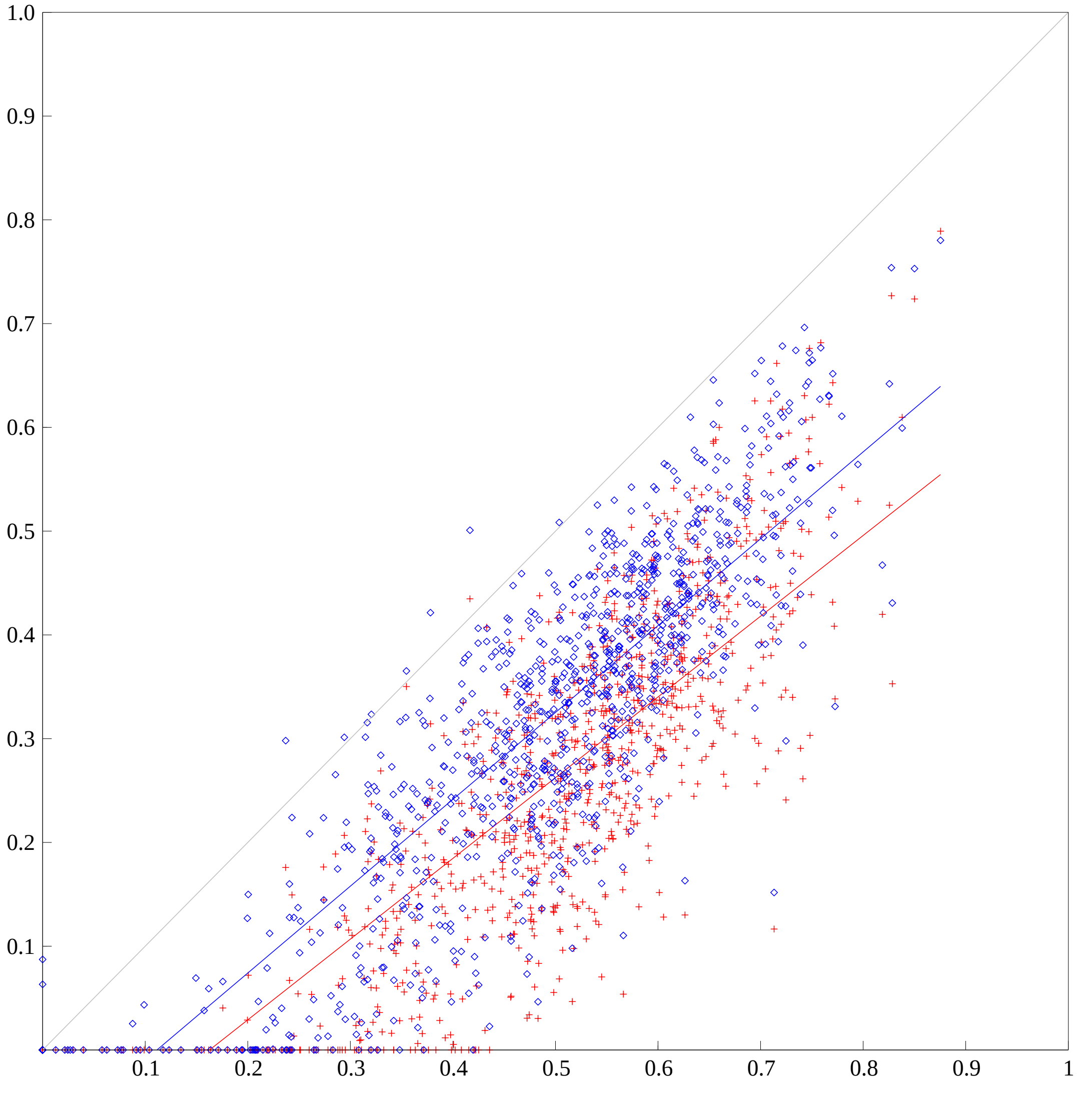
Figure 6: Band spectral estimation and seasonally adjusted series: one typical replication from the Monte Carlo experiment.



by fitting the RWpN model to the seasonally adjusted series ($\tilde{\theta}_{SA}$) versus $\tilde{\theta}$, the standard MLEs of the same parameter obtained from fitting the original trend plus seasonal plus irregular model to the simulated raw series. The solid line is the least squares regression line. We superimpose the scatterplot of the BS estimates (also computed on the seasonally adjusted series) with $k = 8$ (i.e. suppressing all the frequencies in a neighbourhood ± 0.4 around the seasonal frequencies - the conclusions are unchanged if we increase k), denoted $\tilde{\theta}_{BS}$, versus $\tilde{\theta}$; the regression line is now plotted as a dashed line. The plot shows that θ is underestimated and pushed towards its lower limit. In fact, most of the points lie below the 1-1 line ($\tilde{\theta}_{SA} = \tilde{\theta}$) and there is a concentration of points along the line ($\tilde{\theta}$,0). The underestimation is less severe for the BS estimates; the regression line is moved upwards in the direction of the 1-1 line ($\tilde{\theta}_{BS} = \tilde{\theta}$), but there are no ways of improving the situations by increasing k. We also experimented with other weighting patterns, such as pure truncation at a particular cutoff, but these did not prove more effective.

We conclude that BS estimation is helpful for diagnosing and partially correcting the upward bias in the trend variance estimate. BS eliminates the frequencies that are most affected by the

Figure 7: Plot of standard and BS maximum likelihood estimates of the moving average parameter θ, denoted $\tilde{\theta}_{SA}$ and $\tilde{\theta}_{BS}$ (seasonally adjusted series) versus $\tilde{\theta}$, the standard MLE computed on the original simulated series.



adjustment. Nevertheless, for certain parameter combinations the type of overadjustment illustrated by the last panel calls for more clever ways of reviving by reweighting the high frequency components of the spectrum that falls between the seasonal frequencies.

7 Band spectral estimation of approximate models

The previous sections have dealt with the case when a parametric model is thought to be correctly specified for a series y_t, for which only a transformed version, y_t^*, is available. The typical transformations are moving average smoothing, disaggregation of temporal aggregates and seasonal adjustment.

We now address a different situation, according to which the RWpN model is thought only as an approximation. As a matter of fact, it features a stripped to the bone parametric representation for the components, postulating a naïve separation of the short run from the long run fluctuations, such that the trend and the cycle are formulated in terms of the simplest linear and Gaussian nonstationary and stationary processes, respectively.

Due to its simplicity, the RWpN is usually misspecified for economic time series such as gross domestic product (GDP), e.g. because the short run dynamics may display richer structure. Our objective is to extract a random walk trend, so the model can be regarded as only partially specified.

Rather than fitting a more general model, e.g. specifying a stationary ARMA model for ϵ_t, we focus on the strategy of adopting for the RWpN an alternative estimation method so as to enhance the separation of the short run from the long run.

In the time series literature considerable attention has been attracted by the issue of determining alternative estimation criteria that could guarantee efficiency in prediction and signal extraction from a model that would otherwise be rejected from the standard model building methodology. In the next subsection we explore the connections with multistep estimation.

The band spectral approach illustrated in this paper implements explicitly the idea of weighting differently the contribution of each Fourier frequency to the likelihood of the model, thereby enhancing the separability of the trend from the cycle.

We illustrate the role of BS estimation with respect to the problem of extracting the trend from the logarithms of the Italian and the U.S. quarterly GDP, using the RWpN as an approximation to the true model of economic fluctuations. The first series is available from 1970.1 to 2005.1 (Source OECD Statistical Compendium), and the second from 1947.1 to 2005.2 (Source: Bureau of Economic Analysis). The logarithms of the series are displayed in Figure 8.

The maximiser of (2) is $\tilde{\theta} = 0$ in both cases (equivalently $\tilde{\sigma}_\epsilon^2 = 0$), implying that the trend is coincident with the observations; the usual diagnostics highlight the presence of misspecification. The empirical spectrum of Δy_t, displayed in the right plots of figure 9, only for the frequency range $[0, \pi]$ due to its symmetry around π, is interpolated by a constant spectrum; the resulting trend extraction filter uses only the current observation with unit weight.

If we ignore the high frequencies we get local likelihood estimates that move away from zero, imply smoother trends. This fact is illustrated by figure 9: the plots on the left hand side display the estimated θ values for cutoff frequencies in the range $\pi/12$ (corresponding to a period of 6 years) and π (which amounts to considering the standard likelihood (2)). Those on the right hand side display, along with the empirical spectrum of Δy_t, the parametric spectral density implied by the RWpN, that has been fitted using the empirical spectrum up to the cutoff frequency $\omega_c = \pi/6$ (corresponding to a period of 3 years).

In the Italian GDP case, when $.5 < \omega_c < 1$, $\tilde{\theta}$ moves away from zero. To get an idea of the level of smoothing implied by the local likelihood estimate using $\omega_c = \pi/6$, one should refer to figure 8, which displays the smoothed estimates of the trend and the irregular corresponding to the maximiser of (3) using the above cutoff. For the U.S. case $\tilde{\theta}$ is positive and high for low cutoffs, and decreases to zero more gradually than in the Italian case.

7.1 Relationship with multistep estimation

The band spectral approach is closely related to multistep estimation (ME) of the RWpN, also known as adaptive estimation, which is discussed by a rich literature starting from the seminal paper of Cox (1961), among which Tiao and Xu (1993) and Haywood and Tunicliffe Wilson (1997) stand out prominently.

According to ME the parameters are estimated by minimising the variance of the l-step-ahead prediction error, with $l > 1$. The previous references show the merits of ME of the RWpN model for the purpose of forecasting, when the true model is not coincident with the fitted one. The

Figure 8: BS estimation of the RWpN model for the Italian and U.S. quarterly GDP (logarithms). Trend and irregular estimates corresponding the estimated θ values using the cutoff frequency $\pi/6$ (corresponding to a period of 3 years).

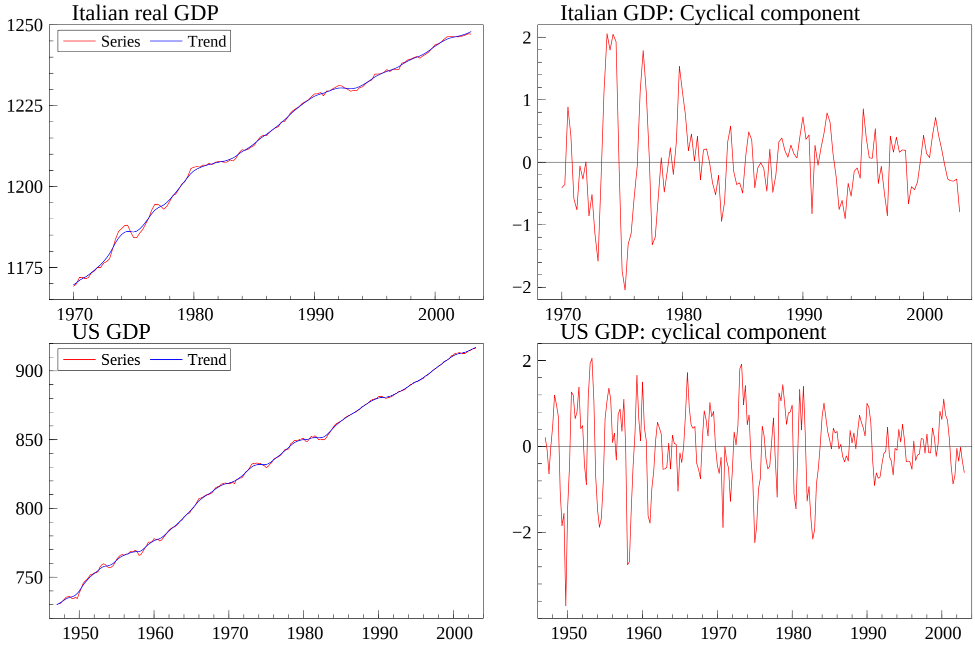
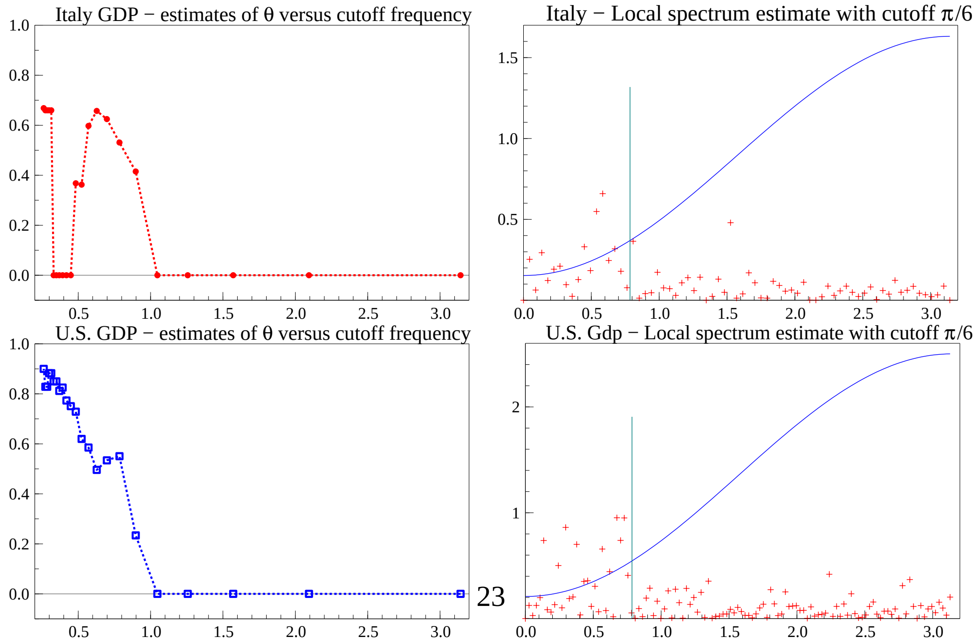


Figure 9: BS estimation of the RWpN model for the Italian and U.S. quarterly GDP (logarithms). Left hand side: estimated θ values for cutoff frequencies in the range $\pi/12$ and π. Right hand side: empirical spectrum of Δy_t, and the parametric spectral fit using $\omega_c = \pi/6$.



efficency stems from the fact that the estimation criterion selects the information that is relevant for predicting the series for medium to long forecast horizons. Proietti (2005) illustrates the role of ME for signal extraction.

The l-step-ahead forecast errors can be written as follows:

$$\nu_t(l) = \frac{n(L)}{1-\theta L}\Delta y_t, \quad n(L) = 1 + (1-\theta)L + \cdots (1-\theta)L^{l-1}.$$

Adopting the Fourier transform and using the approximation in Haywood and Tunicliffe Wilson (1997, p. 242) we can express the ME estimator of θ as

$$\min_{\theta} \left\{ \sum_{j=0}^{T-1} \mathrm{w}_j \frac{2\pi I(\omega_j)}{g(\omega_j)} \right\}, \quad \mathrm{w}_j = \left| n(e^{-\imath\omega_j}) \right|^2.$$

Thus, the weights are provided by the squared gain of the filter $n(L)$. Given the properties of the latter, ME is equivalent to a weighted estimation criterion placing more weight on the low-frequencies.

In multistep estimation the kernel is automatically provided by the forecast function of the approximating model and depends on its parameters; thus, in the RWpN case it depends on θ. In the BS, instead, the kernel is independent of θ.

8 Conclusions

This paper has focused on band spectral estimation of the random walk plus noise model. This estimation strategy proves effective in two situations: the first occurs when the high frequency components in the spectrum have been distorted by prefiltering, aiming at enhancing the smoothness of the series, temporal disaggregation, and seasonal adjustment.

Band spectral estimation can be an effective tool for eliciting a long-run trend from a time series using a simplified model of economic fluctuations. As the illustrations have shown, it conceptualises and operationalises the notion of constructing simple predictors and signal extraction filters that enhance the separation of the long run features of a series, by giving more consideration to how well the model fits the empirical spectrum at the low (long run) frequencies.

More generally, it is a useful device to diagnose whether a proposed parametric model is valid across the entire frequency range.

References

Corbae, D., Ouliaris, S. and Phillips, P.C.B. (2002). Band Spectral Regression with Trending Data, *Econometrica*, 70, 1067-1109.

Cox, D.R. (1961). Prediction by exponentially weighted moving averages and related methods. *Journal of the Royal Statistical Society*, Series B, 23, 414-422.

Christiano, L.J., and Vigfusson, R.J. (2003). Maximum likelihood in the frequency domain: the importance of time-to-plan, *Journal of Monetary Economics*, 50, 789815.

Diebold, F.X., Ohanian, L.E., Berkowitz, J. (1998). Dynamic equilibrium economies: a framework for comparing models and data. *Review of Economic Studies* 65, 433451.

Doornik, J.A. (2001), *Ox. An Object-Oriented Matrix Programming Language*, Timberlake Consultants Press, London.

Durbin J., and Koopman, S.J. (2001). *Time Series Analysis by State Space Methods*, Oxford University Press: New York.

Engle, R. F. (1974). Band Spectrum Regression, *International Economic Review*, 15, 1-11.

Engle, R. F. (1978). Testing Price Equations for Stability across Spectral Frequency Bands, *Econometrica*, 46, 869-881.

Gardner, E. S. (1985). Exponential Smoothing: The State of the Art, *Journal of Forecasting*, 4, 128.

Hannan, E.J. (1969). The Estimation of Mixed Moving Average Autoregressive Systems, *Biometrika*, 56, 579-593.

Harvey, A.C. (1989). *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

Harvey, A.C. and Chung, C.H. (2000). Estimating the underlying change in unemployment in the UK, *Journal of the Royal Statistics Society, Series A*, 163, 303-339.

Harvey, A.C. and Koopman, S.J. (2000). Signal Extraction and the Formulation of Unobserved Components, *The Econometrics Journal*, 3, 84-107.

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Springer Series in Statistics, Springer.

Haywood, G. and Tunnicliffe Wilson, G. (1997), Fitting Time Series Models by Minimizing Multistep-ahead Errors: a Frequency Domain Approach, *Journal of the Royal Statistical Society, Series B*, 59, 237-254.

Jaeger, A. (1992). Does Consumption Take a Random Walk? Some Evidence from Macroeconomic Forecasting Data, *The Review of Economics and Statistics*, 74, 607-614

Koopman, S.J., Shephard, N., and Doornik, J.A. (1999). Statistical algorithms for models in state space using SsfPack 2.2, *Econometrics Journal*, 2, 113-166.

Muth, J.F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, 55, 299-306.

Nerlove, M.L., Grether, D.M. and Carvalho, J.L. (1995), *Analysis of Economic Time Series: A Synthesis*, Revised edition, Academic Press Inc., New York.

Pollock, D.S.G. (2006a). Estimation with Band-Limited data, *Mimeo*.

Pollock, D.S.G. (2006b). Investigating economic trends and cycles, *Mimeo*.

Proietti T. (2005). Forecasting and Signal Extraction with Misspecified Models, *Journal of Forecasting*, 24, 539556.

Proietti T. (2006). Temporal Disaggregation by State Space Methods: Dynamic Regression Methods Revisited. *Econometrics Journal*, 9, 357372.

Robinson, P.M. (1977). The construction and estimation of continuous time models and discrete approximations in econometrics. *Journal of Econometrics*, 6, 173-197.

Shephard, N.G. (1993). Maximum likelihood estimation of regression models with stochastic trend components, *Journal of the American Statistical Association*, 84, 590-95.

Tiao, G. C., and Xu, D. (1993). Robustness of Maximum Likelihood Estimates for Multi-Step Predictions: The Exponential Smoothing Case. *Biometrika*, 80, 623-641.

Thomson, P.J. (1986). Band-Limited spectral estimation of autoregressive-moving-average processes. *Journal of Applied Probability*, 23A 'Essays in Time Series and Allied Processes. Papers in honour of E.J. Hannan' (ed. J. Gani and M.B. Priestley), 143-155.

Watson, M., (1993). Measures of fit for calibrated models. *Journal of Political Economy*, 101, 10111041.

Whittle P. (1983) *Prediction and Regulation by Linear Least Squares Methods*, Second edition. Basil Blackwell, Oxford.