March 16, 1998

The Beveridge–Nelson Decomposition: a Different Perspective With New Results

Víctor Gómez

Ministerio de Economía y Hacienda, Madrid

and

Jörg Breitung

Humboldt Universität zu Berlin
Institut für Statistik und Ökonometrie

Abstract

We show in the paper that the decomposition proposed by Beveridge and Nelson (1981) for models that are integrated of order one can be generalized to seasonal ARIMA models by means of a partial fraction decomposition. Two equivalent algorithms are proposed to optimally (in the mean squared sense) compute the estimates of the components in the generalized decomposition. While the first algorithm is very fast and easy to implement, the second can also provide the standard errors of the estimated components. The properties of the implied filters are investigated and compared with those obtained using the model–based TRAMO/SEATS software package. The alternative methods are applied to the German unemployment series.

A part of the research for this paper was carried out within Sonderforschungsbereich 373 at the Humboldt University Berlin and was printed using funds made available by the Deutsche Forschungsgemeinschaft.

1 Introduction

In recent years, several model–based seasonal adjustment procedures have been suggested to overcome the ad–hoc character of widely used procedures based on moving average filters, like Census X-11. Due to the influence of the Box and Jenkins' (1970) methodology, most of these model–based approaches postulated components following Arima models. For example, Box et al. (1978), Nerlove et al. (1979), Harvey and Todd (1983) and Maravall and Pierce (1987) adopted an unobserved Arima components framework with orthogonal components. Since the development of the software packages Tramo/Seats (Gómez and Maravall, 1997) and Stamp (Harvey, 1984), this type of approach has become increasingly popular in practice[1].

Another approach which is based on an Arima framework is the one advocated by Beveridge and Nelson (1981) for nonseasonal series which are integrated of order one. This approach has been extended to some seasonal models by Newbold and Vougas (1995). However, to the best of our knowledge, a complete solution to the Beveridge–Nelson type of decomposition for the general case of finite nonstationary seasonal series, integrated of any order, has not been given in the literature. In this paper, we give one such general solution, together with two efficient algorithms which allow for the computation of the estimates of the components and, if desired, also their mean squared errors.

The proposed solution is based on a partial fraction decomposition of the Arima model followed by the series, since, as we show in the paper, that is precisely what the decomposition proposed by Beveridge and Nelson (1981) amounts to in the case of a nonseasonal series which is integrated of order one.

[1]The Tramo/Seats software has the potential to compete with the new Census X-12 program, since it has the capabilities of automatic model identification, automatic outlier treatment, pre–testing of Trading Day and Easter effects, etc. Besides, it can be efficiently used for routine application to a large number of series, as is done, for example, at Eurostat

The two algorithms proposed to estimate the components are, first the Kalman filter and smoother, appropriately initialized because the series are nonstationary, and, second a kind of G. Tunnicliffe Wilson's algorithm like the one proposed in Burman (1980). We show in the paper that both algorithms give identical results. The first algorithm, however, is the only one that can give the standard errors. Besides, it can also be used in the cases in which there are fixed effects or missing observations in the model, where the second algorithm cannot be applied.

Since the proposed Beveridge–Nelson decomposition, henceforth referred to as BND, starts with an ARIMA model for the series at hand and from that model constructs models for the components according to a certain rule, the question naturally arises as to whether the results obtained with this procedure will be similar to the results obtained with the TRAMO/SEATS procedure. The latter procedure obtains models for the components from the model followed by the series by imposing the so–called canonical decomposition. See Maravall and Pierce (1987).

There is one fundamental difference, however, between both procedures. In the TRAMO/SEATS approach, hereafter referred to as TSA, the series is assumed to be the sum of a certain number of orthogonal components. Since the components are unobserved, they are estimated by means of the (finite version of the) Wiener–Kolmogorov filters or, what amounts to the same thing, the Kalman filter and smoother. In the case of the BND, the models for the components are not orthogonal because all the innovations of these models coincide with the innovations of the model for the series. This implies that, if we knew the whole past of the series, the components of the BND would be in fact observed. Since the observed series is finite, we will have to estimate the components of the BND by projecting the unknown past values of the series onto the finite sample. Therefore, the relevant comparison is between the minimum mean squared error (MMSE) estimators of the (correlated) components of the BND and the MMSE estimators of the (orthogonal) components of the TSA.

We compare in the paper the filters obtained with the BND with those obtained by the TSA for some of the more usual models in practice. The proposed methodology is applied to the series of German unemployment and the results are compared with the ones obtained applying the TSA.

The paper is structured as follows. In Section 2, the decomposition originally proposed by Beveridge and Nelson (1981) is reviewed and the result is established that this decomposition can be obtained by means of a partial fraction decomposition. Also in this Section, and based on the previous result, a BND is proposed for general ARIMA models. The two algorithms to estimate the components in the proposed BND are described in Section 3 and their equivalence is proved. In Section 4, the properties of the filters for the components obtained with the BND are studied and compared to those of the corresponding filters obtained with the TSA. In Section 5, both approaches are applied to the German unemployment series. Section 6 summarizes the conclusions.

2 A General Framework for the Beveridge–Nelson Decomposition

Beveridge and Nelson (1981) proposed a decomposition for ARIMA$(p, 1, q)$ models which was further investigated by Cuddington and Winters (1987), Miller (1987) and Newbold (1990). Suppose $\{z_t\}$ is an $I(1)$ process such that ∇z_t has the Wold decomposition $\nabla z_t = \Psi(B)a_t$, where B is the backshift operator, $Ba_t = a_{t-1}$, and $\nabla = 1-B$. Then, according to Beveridge and Nelson (1981), z_t can be expressed as the sum of a permanent p_t and a transitory c_t component, where p_t is defined as the sum of the current observed value z_t and all forecastable future changes in the series. It was shown by these authors that the previous definition implies that the permanent component follows the model $\nabla p_t = \Psi(1)a_t$ and the transitory component is given by $c_t = \Psi^*(B)a_t$, where $\Psi^*(B)$ satisfies $(1-B)\Psi^*(B) = \Psi(B) - \Psi(1)$.

It is a remarkable fact that the BND can also be obtained by means of a partial fraction expansion of the rational lag function $\Psi(B)$. Before proving this result, we summarize the partial fraction expansion of a rational function in the following lemma.

Lemma 1 *Let $a(x)$ and $b_j(x)$ be polynomials of degree $n \geq 0$ and $m_j \geq 1$ for $j = 1, \ldots, K$. Then, the partial fraction expansion yields the unique decomposition*

$$\frac{a(x)}{\prod_{j=1}^{K} b_j(x)} = c(x) + \sum_{j=1}^{K} \frac{d_j(x)}{b_j(x)}, \tag{1}$$

where $c(x)$ and $d_j(x)$ are polynomials of degree $n^* = max\{0, n - \bar{m}\}$ and $m_j^* = m_j - 1$, respectively, and $\bar{m} = \sum_{j=1}^{K} m_j$. It is understood that $c(x) = 0$ if $n < \bar{m}$.

A proof of the lemma can be found, for example, in Van der Waerden (1970). The polynomials $c(x)$ and $d_j(x)$ $(j = 1, \ldots, K)$ can be determined by multiplying (1) with $\sum b_j$ and comparing the coefficients of the resulting polynomials. This yields a system of linear equations which is solved for the coefficients of $c(x)$ and $d_j(x)$. Alternatively, one can successively multiply (1) by the different factors of the b_j polynomials while setting the variable x equal to the roots of these factors. In this way, one can sequentially obtain the unknown coefficients.

To see that the BND can be obtained by means of a partial fraction expansion, suppose first that $\{z_t\}$ follows the model $\phi(B)\nabla z_t = \theta(B)a_t$, where the polynomial $\phi(B)$ is of degree p and has all its roots outside the unit circle and the polynomial $\theta(B)$ is of degree q. Then, write the partial fraction expansion

$$\frac{\theta(x)}{\phi(x)(1-x)} = \gamma(x) + \frac{k}{1-x} + \frac{\alpha(x)}{\phi(x)}, \tag{2}$$

where $\gamma(x)$ is a polynomial of degree $max\{0, q - p - 1\}$, k is a constant and $\alpha(x)$ is a polynomial of degree $max\{0, p - 1\}$.

Multiplying (2) by $1 - x$ and letting $x = 1$ yields $k = \Psi(1)$, where $\Psi(x) = \theta(x)/\phi(x)$ is, as before, the expression that gives the weights in the Wold decomposition of ∇z_t. Defining $\gamma(x) + \alpha(x)/\phi(x) = \eta(x)/\phi(x)$, it is not difficult to verify that $\Psi(x) = \Psi(1) + (1-x)\eta(x)/\phi(x)$, so that, with the previous notation, $\Psi^*(x) = \eta(x)/\phi(x)$. Letting $\nabla p_t = ka_t$ and $\phi(B)c_t = \eta(B)a_t$, the result is proved. Note that the degree of $\eta(B)$ is $max\{p-1, q-1\}$, which coincides with the result of Newbold and Vougas (1995).

The original BND allowed for a constant in the model. This is easily incorporated into the present context, since, in the partial fraction expansion we would obtain the extra term $\mu/(1-x)$, which would be assigned to the trend component. This follows from the fact that now the model for the series is $\phi(B)(\nabla z_t - \mu) = \theta(B)a_t$. The model for the trend would be $\nabla p_t = \mu + \Psi(1)a_t$.

The previous result can be generalized to the multiplicative seasonal ARIMA model of the type

$$\phi(B)\Phi(B^n)(\nabla^d\nabla_n^D z_t - \mu) = \theta(B)\Theta(B^n)a_t, \qquad (3)$$

where, μ is the mean of the differenced series, n is the number of seasons, $d = 0,1,2$, $D = 0,1$, $\nabla = 1 - B$ is a regular difference and $\nabla_n = 1 - B^n$ is a seasonal difference. Instead of z_t, it may be necessary to use $log(z_t)$, or some other transformation, to stabilize the variance of the series. If p and P are the degrees of the autoregressive polynomials $\phi(B)$ and $\Phi(B)$, and q and Q those of the moving average polynomials $\theta(B)$ and $\Theta(B)$, model (3) is denoted as a multiplicative $(p,d,q)(P,D,Q)_n$ model.

If we try to apply the original Beveridge and Nelson's idea, which is based on the forecast function of the series, to decompose z_t in (3) into a trend, a seasonal and an irregular component, the task seems formidable; see, for example, the paper by Newbold and Vougas (1995). However, if we make use of the partial fraction expansion, we immediately obtain a unique decomposition which makes sense. To this end, first define the polynomials $\phi^*(B) = \phi(B)\Phi(B^n)$, $\Delta(B) = \nabla^d\nabla_n^D$ and $\theta^*(B) = \theta(B)\Theta(B^n)$ and suppose for simplicity that there is no mean in (3). Then, consider the partial fraction decomposition

$$\frac{\theta^*(x)}{\phi^*(x)\Delta(x)} = \gamma(x) + \frac{\alpha_p(x)}{(1-x)^{d+D}} + \frac{\alpha_s(x)}{S(x)} + \frac{\alpha_c(x)}{\phi^*(x)}, \tag{4}$$

where $S(x) = 1 + x + \cdots + x^{n-1}$ and the third term on the right of the previous expression exists only if $D > 0$. Note that we have used in (4) the fact that $\nabla_n = (1-B)S(B)$. The degrees of the $\gamma(x)$, $\alpha_p(x)$, $\alpha_s(x)$ and $\alpha_c(x)$ polynomials in (4) are, respectively, $max\{0, q^* - p^* - d^*\}$, $d^* - 1$, $n - 2$ and $p^* - 1$, where $p^*{=}p + P$, $q^* = q + Q$ and $d^* = d + D$.

We could further decompose $S(B)$ in (4) into its different seasonal factors, what would give rise in turn to subcomponents associated with the different seasonal frequencies. However, in order to simplify matters, we will consider in this paper a unique seasonal component, which will be given by the decomposition (4).

The assignment of the terms in (4) to the different components is linked with the roots of the autoregressive polynomials in (3). As for the unit roots, it is clear that the factor $(1-x)^{d^*}$ should be assigned to the trend component p_t, since it corresponds to an infinite peak in the pseudospectrum of the series at the zero frequency. On the other hand, given that all roots of the polynomial $S(x)$ correspond to infinite peaks in the pseudospectrum at the seasonal frequencies, the factor $S(x)$ should be assigned to the seasonal component s_t.

As regards the roots of the autoregressive polynomial $\phi(x)\Phi(x^n)$, the situation is not so clear–cut and the assignment is more subjective. For simplicity, we will consider in what follows only a third component, which will be referred to as "stationary component" c_t. All roots of $\phi(x)\Phi(x^n)$ will be assigned to this stationary component, which, therefore, may include a cyclical component and stationary trend and seasonal components.

Based on the previous considerations, the decomposition $z_t = p_t + s_t + c_t$ is proposed where the trend p_t, seasonal s_t and stationary c_t components are given, respectively, by

$$\nabla^{d^*} p_t = \alpha_p(B)a_t, \quad S(B)s_t = \alpha_s(B)a_t, \quad \phi^*(B)c_t = \eta(B)a_t, \tag{5}$$

where $\eta(x) = \gamma(x)\phi^*(x) + \alpha_c(x)$. We can express the trend, say, in terms of the original series z_t by replacing a_t in (5) by the expression $[\phi^*(B)\Delta(B)/\theta^*(B)]\, z_t$, obtained from (3). This yields

$$p_t = \frac{\alpha_p(B)\phi^*(B)S(B)}{\theta^*(B)} z_t. \tag{6}$$

Therefore, the trend p_t is the result of applying the one–sided filter $H_p(B) = \alpha_p(B)\phi^*(B)S(B)/\theta^*(B)$ to the series z_t. If the roots of $\theta^*(x)$ are all outside the unit circle or, what is the same thing, the model (3) is invertible, we can express (6) as an infinite sum of present and past values of the process $\{z_t\}$, $p_t = \sum_{j=0}^{\infty} \nu_j z_{t-j}$, where $H_p(B) = \sum_{j=0}^{\infty} \nu_j B^j$. Since in practice we only know a finite series $z = (z_1, \ldots, z_N)'$, we will have to estimate the unknown z_t in the previous expression with backcasts.

Proceeding similarly, we find for the other two components of the series that $s_t = H_s(B)z_t$ and $c_t = H_c(B)z_t$, where $H_s = \phi^*(B)\alpha_s(B)\nabla^{d^*}/\theta^*(B)$ and $H_c = \eta(B)\Delta(B)/\theta^*(B)$.

A couple of examples will serve to clarify matters. Let the series follow the model $\phi(B)\nabla^2 z_t = \theta(B)a_t$ and let, as before, $\Psi(B) = \theta(B)/\phi(B)$. Then, (4) becomes

$$\frac{\theta(x)}{\phi(x)(1-x)^2} = \gamma(x) + \frac{k_1 + k_2 x}{(1-x)^2} + \frac{\alpha_c(x)}{\phi(x)}, \tag{7}$$

where k_1 and k_2 are constants. Letting $\eta(x) = \gamma(x)\phi(x) + \alpha_c(x)$ and multiplying (7) by $(1-x)^2$ yields

$$\Psi(x) = k_1 + k_2 x + (1-x)^2 \frac{\eta(x)}{\phi(x)}. \tag{8}$$

Letting $x = 1$, it is obtained that $\Psi(1) = k_1 + k_2$. Differentiating (8) and letting again $x = 1$, we get $\Psi'(1) = k_2$, where $\Psi'(1)$ is the derivative of $\Psi(x)$ evaluated at $x = 1$. From this, it is obtained that $k_1 = \Psi(1) - \Psi'(1)$ and $k_2 = \Psi'(1)$. The trend p_t follows the model $\nabla^2 p_t = (k_1 + k_2 B)a_t$, which coincides with the model obtained by Newbold and Vougas (1995).

As a second example, consider the model $\phi(B)\nabla_n z_t = \theta(B)a_t$ and let again $\Psi(B) = \theta(B)/\phi(B)$. Then, (4) becomes

$$\frac{\theta(x)}{\phi(x)(1-x^n)} = \gamma(x) + \frac{k}{1-x} + \frac{\alpha_s(x)}{S(x)} + \frac{\alpha_c(x)}{\phi(x)}, \tag{9}$$

where k is a constant. Multiplying (9) by $1 - x$ and letting $x = 1$ yields $k = \Psi(1)/s$. The model for the trend p_t is $\nabla p_t = k a_t$, which again coincides with the one obtained by Newbold and Vougas (1995).

The same argument shows that the model for the trend given by the proposed procedure for the model $\phi(B)\nabla\nabla_n z_t = \theta(B)a_t$ also coincides with the one obtained by Newbold and Vougas (1995). Therefore, the proposed procedure is completely general and encompasses all models for which a BND has been given so far.

3 Two Algorithms to Estimate the Beveridge–Nelson Decomposition

In this section we will describe two algorithms to estimate the components in the proposed BND. The first algorithm is very simple and is based on the algorithm proposed by G. Tunnicliffe Wilson in Burman (1980). The second algorithm consists of the Kalman filter plus a smoothing algorithm, with a proper initialization for the Kalman filter because the series is nonstationary. Both algorithms will be proved to be equivalent.

All nonstationary series considered in the proposed BND, that is, the original series z_t, the trend p_t and the seasonal s_t components, are assumed to start at some finite time in the past. They are supposed to be generated like in Bell (1984), as linear combinations of some starting values and elements of the differenced processes. Note however that, in the present context, all series have the same innovations.

In the rest of the section we suppose that the series follows the general Arima model (3), where, for simplicity, it is assumed that $\mu = 0$.

The Backcasting Algorithm

Since all three components can be expressed as the output of one–sided filters applied to the process $\{z_t\}$, suppose that we want to estimate the filtered series $y_t = H(B)z_t$, where $H(B)$ is the ARMA filter $H(B) = \rho(B)/\theta^*(B)$ and y_t is any of the three components. By the results of the previous section, $\rho(x)$ is $\alpha_p(x)\phi^*(x)S(x)$ in the case of the trend, $\alpha_s(x)\phi^*(x)(1-x)^{d^*}$ in the case of the seasonal, and $\eta(x)(1-x)^d(1-x^n)^D$ in the case of the stationary component. The proposed algorithm will avoid the computation of an infinite number of backcasts. In fact, it will be necessary to computute only a small number of backcasts.

Given that the series z_t also follows the backward model $\phi^*(F)\Delta z_t = \theta^*(F)v_t$, where F is the forward operator, $Fz_t = z_{t+1}$, projecting onto the finite sample $z = (z_1, \ldots, z_N)'$ implies $\phi^*(F)\Delta z_t = 0$, $t \leq -q^*$. Then, letting r be the degree of $\rho(x)$, the first algorithm is

> Solve the system
>
> $$\begin{aligned} \theta^*(B)y_t &= \rho(B)z_t \quad t = -q^*+1, \ldots, p^*-q^* \\ \phi^*(F)\Delta y_t &= 0 \quad\quad\quad t = -2q^*+1, \ldots, -q^* \end{aligned}$$
>
> where $q^* + r$ backcasts are needed: $\hat{z}_{-q^*-r+1}, \ldots, \hat{z}_0$.
> For $t = p^* - q^* + 1, \ldots, N$, obtain y_t from the recursion $\theta^*(B)y_t = \rho(B)z_t$.

In order to obtain the backcasts needed in the previous algorithm, we can use the Kalman filter like in Gómez and Maravall (1994) with the reversed series.

To illustrate, consider the very simple example $\nabla_2 z_t = a_t$. Then, the partial fraction decomposition

$$\frac{1}{1-x^2} = \frac{c_1}{1-x} + \frac{c_2}{1+x}$$

yields $c_1 = c_2 = 1/2$, so that the trend p_t follows the model $\nabla p_t = (1/2)a_t$ and the seasonal s_t follows the model $S(B)s_t = (1/2)a_t$, where $S(B) = 1+B$.

There is no stationary component in this case. It is straightforward to verify that $p_t = (1/2)(1+B)z_t$ and $s_t = (1/2)(1-B)z_t$. In order to estimate the trend, only one backcast $\hat{z}_0$ is needed, which is easily seen to be $\hat{z}_0 = z_2$. This implies $\hat{p}_1 = (z_1 + z_2)/2$.

The Kalman Filter Algorithm

This algorithm consists of an augmented Kalman filter plus an augmented smoothing algorithm. The state space representation we will use is based on that of Gómez and Maravall (1994), which in turn is an extension to nonstationary models of the representation proposed by Akaike (1974) for ARMA models.

The state space representation of the proposed BND is obtained from the state space representation of each component. Following Gómez and Maravall (1994), the state space representation of the trend, for example, can be obtained as follows. Let m be the degree of $\alpha_p(x)$ in (5), let $\nabla^{d^*} = 1 + \phi_1^p B + \cdots + \phi_{d^*}^p B^{d^*}$, $r^p = max\{d^*, m+1\}$ and define $\phi_i^p = 0$ when $i > d^*$. Then, the state space representation of the trend p_t is

$$p_t = H_p' x_t^p \tag{10}$$

$$x_{t+1}^p = F_p x_t^p + G_p a_{t+1}, \tag{11}$$

where

$$F_p = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\phi_{r^p}^p & -\phi_{r^p-1}^p & -\phi_{r^p-2}^p & \dots & -\phi_1^p \end{bmatrix},$$

$x_t^p = (p_t, p_{t+1|t}, \dots, p_{t+r^p-1|t})'$, $H_p = (1, 0, \dots, 0)'$, $G_p = (1, \psi_1^p, \dots \psi_{r^p-1}^p)'$ and the ψ_i^p weights are the coefficients obtained from $\psi^p(B) = \alpha_p(B)/\nabla^{d^*} = \sum_{i=0}^{\infty} \psi_i^p B^i$. The elements of the state vector are defined as $p_{t+i|t} = p_{t+i} - \psi_0^p a_{t+i} - \cdots - \psi_{i-1}^p a_{t+1}$, $i = 1, \dots, r^p - 1$. They are the predictors of p_{t+i} based on the semi-infinite sample $\{p_j : j \le t\}$. The state space representations for the seasonal s_t and the stationary c_t components are defined by replacing in (10) and (11) p with s and c, respectively, where the matrices F_s, F_c, G_s, G_c, H_s and H_c and the vectors x_t^s and x_t^c are defined analogously to F_p, G_p, H_p and x_t^p.

The state space representation of the proposed BND is defined by

$$z_t = H'x_t \tag{12}$$

$$x_{t+1} = Fx_t + Ga_{t+1}, \tag{13}$$

where $F = \text{diag}[F_p, F_s, F_c]$ (a block diagonal matrix), $G = (G'_p, G'_s, G'_c)'$, $H' = (H'_p, H'_s, H'_c)$ and $x_t = (x_t^{p'}, x_t^{s'}, x_t^{c'})'$.

In order to obtain initial conditions for the augmented Kalman filter, to be applied to (12) and (13), we will proceed like in Bell (1984) with the two nonstationary components p_t and s_t. That is, we will generate these components as linear combinations of some starting values and elements of the differenced processes, $u_t^p = \nabla^{d^*} p_t$ and $u_t^s = S(B)s_t$. We will illustrate the process with the trend component p_t. Let the starting values be $\delta^p = (p_{1-d^*}, \ldots, p_0)'$. Then, following Bell (1984), the p_t can be generated from $p_t = A_t^{p'}\delta^p + \sum_{i=0}^{t-1} \xi_i^p u_{t-i}^p$, where $t > 0$, $1/\nabla^{d^*} = \sum_{i=0}^{\infty} \xi_i^p B^i$ and the $A_t^p = (A_{1,t}^p, \ldots, A_{d^*,t}^p)'$ can be recursively generated from

$$\begin{aligned} A_t^p &= (0, \ldots, 1, \ldots, 0), & t &= 1 - d^*, \ldots, 0 \\ A_t^p &= -\phi_1^p A_{t-1}^p - \cdots - \phi_{d^*}^p A_{t-d^*}^p, & t &> 0, \end{aligned}$$

where for $t = 1 - d^*, \ldots, 0$ the one is in the $(t + d^*)$–th position. Note that we have used $(p_{1-d^*}, \ldots, p_0)'$ as starting values instead of the starting values$(p_1, \ldots, p_{d^*})'$ used by Bell (1984). This is immaterial for the theoretical development and is done to facilitate the initialization of the Kalman filter algorithm.

Like in Gómez and Maravall (1994), p. 615, it can be shown that the initial state vector x_1^p in (10) and (11) verifies $x_1^p = A^p\delta^p + \Xi^p U^p$, where $A^p = [A_1^p, \ldots, A_{r^p}^p]'$, Ξ^p is the lower triangular matrix with rows the vectors $(\xi_{j-1}^p, \xi_{j-2}^p, \ldots, 1, 0, \ldots, 0)$, $j = 1, \ldots, r^p$, $U^p = (u_1^p, u_{2|1}^p, \ldots, u_{r^p|1}^p)'$ and $u_{i|1}^p$

$= \mathrm{E}(u_i^p | u_t^p : t \leq 1)$, $i > 1$. A similar argument would lead to an initial state vector $x_1^s = A^s \delta^s + \Xi^s U^s$ for the seasonal component s_t, where the starting values are $\delta^s = (s_{n-1}, \ldots, s_0)'$ and the matrices A^s, Ξ^s and U^s are defined analogously to A^p, Ξ^p and U^p. Since the stationary component c_t is, by definition, stationary no special assumptions need to be made with respect to its initial state vector x_1^c.

Then, the initial state vector for (12) and (13) is given by $x_1 = A\delta + \Xi U$, where $A = [\bar{A}', 0']'$, $\bar{A} = \mathrm{diag}[A^p, A^s]$, $\delta = (\delta^{p'}, \delta^{s'})'$, $\Xi = \mathrm{diag}[\Xi^p, \Xi^s, I]$, and $U = (U^{p'}, U^{s'}, x_1^{c'})'$.

The augmented Kalman filter and smoothing algorithms that we will use are based on the results of De Jong (1991). A review of De Jong's results, together with a detailed description of the proposed algorithms can be found in the appendix.

Continuing with the example of this section, the state space representation is given by (12) and (13), where $x_t = (p_t, s_t)'$, $H' = (1, 1)$, $G = (1, 1)/2$ and

$$F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The initial state is $x_1 = F\delta + Ga_1$, where $\delta = (p_0, s_0)'$. The initial conditions for the augmented Kalman filter are $(\hat{x}_{1|0}, \hat{X}_{1|0}) = (0, -F)$ and $\Sigma_{1|0} = GG'$.

We finish this section by proving that the two proposed algorithms yield identical estimates of the components. The result is contained in the following theorem.

Theorem 1 *The two algorithms described in this section to estimate the components of the proposed BND yield identical results.*

PROOF. Suppose that we want to use the second algorithm described in this section to estimate all components based on $z = (z_1, \ldots, z_N)'$, the observed series. To this end, the augmented Kalman filter is first applied, followed by the QR algorithm, to obtain the GLS estimator $\hat{\delta}$. Then, we apply the augmented fixed point smoother. Denote this estimator by $\mathrm{E}(y_t | z, \hat{\delta})$, where y_t is any of the three components.

Suppose now that $\eta = (z_{1-d}, \ldots, z_0)'$ is used instead of δ to model uncertainty and that the augmented Kalman filter and smoother are used again to estimate y_t based on z. Denote this estimator by $\mathrm{E}(y_t|z, \hat{\eta})$. Then, by result 1 and examples 1 and 2 of Bell and Hillmer (1991), the transformation approach estimates of y_t using δ and η coincide. By theorem 5.2 of Ansley and Kohn (1985), these last two estimators also coincide with the diffuse estimators $\mathrm{E}(y_t|z, \hat{\delta})$ and $\mathrm{E}(y_t|z, \hat{\eta})$. Note that examples 1 and 2 continue to be valid in the present context, although the components are now correlated, because $\eta = M\delta + u$, where M is nonsingular and u is a stationary vector. It is only the stationary vector u that changes with respect to the context with orthogonal components.

Make assumption A of Bell (1984) and suppose that η instead of $z_* = (z_1, \ldots, z_d)'$ is used to generate the series and assume that the semi–infinite realization $\{\ldots, z_{-1}, z_0, z_1, \ldots, z_N\}$ is known. Then, using the notation and results of this section, the component y_t satisfies the difference equation $\theta^*(B)y_t = \rho(B)z_t$. Projecting first both sides of this equation onto the space generated by $\{\eta, z_1, \ldots, z_N\}$, where η is considered fixed, and then projecting again onto the space generated by $\{z_1, \ldots, z_N\}$, yields $\theta^*(B)\hat{y}_t = \rho(B)\hat{z}_t$, where $\hat{z}_t = z_t$ for $t = 1, \ldots, N$ and is a backcast based on z otherwise, and $\hat{y}_t = \mathrm{E}(y_t|z, \hat{\eta})$, the estimator mentioned above. The backcasts can be obtained using the reversed series and an appropriate state space representation for the original series. Since $\mathrm{E}(y_t|z, \hat{\delta}) = \mathrm{E}(y_t|z, \hat{\eta})$, the proof is complete.

As mentioned earlier, the first algorithm can only give the estimates of the components, whereas the second one also gives the standard errors and can be used in a more general context, like, for example, when there are missing observations.

4 The Beveridge–Nelson Decomposition in the Frequency Domain

From the results of the previous sections, it follows that the components of the BND are obtained by applying certain one–sided filters to the original series. More specifically, the trend p_t and seasonal s_t components are obtained by applying the filters H_p and H_s defined immediately after (6). The stationary component c_t can be obtained simply as $c_t = z_t - p_t - s_t$.

To investigate the effects of these filters in the frequency domain, we first discuss a trend model without seasonal component, as in the original work of Beveridge and Nelson (1991). Assume, for simplicity, that the series is generated by the model $\nabla z_t = a_t + \theta a_{t-1}$. Then, as we saw in Section 2, the model for the trend is $\nabla p_t = k a_t$, where $k = 1+\theta$. This implies $p_t = H_p(B) z_t$, where $H_p(B) = [(1+\theta)/(1+\theta B)]\, z_t$. The squared gain $G_p^2(x) = |H_p(e^{-ix})|^2$ results as

$$G_p^2(x) = \frac{(1+\theta)^2}{1+2\theta cosx + \theta^2}.$$

In Fig. 1a and 1b present the gain and phase delay functions for various values of θ. It can be seen from the graphs that as θ goes to -1 the filter behaves better in the sense that it resembles more what one would expect from a trend filter. However, there is a trade–off between better low–pass gain function and greater phase delay in the frequency band around zero. The greater the phase delay, the more the filtered series will have to be shifted to be in phase with the original series. Note that all values in Fig. 1a and 1b correspond to negative values of θ. For positive values of θ, the trend filters behave very badly, taking values much greater than one (not shown).

Consider now the seasonal airline model for quarterly series $\nabla\nabla_4 z_t = (1+\theta B)(1+\Theta B^4)a_t$. To simplify the exposition, a value of $\Theta = -.6$ has been chosen and kept fixed while the parameter θ takes several negative values. The gain and phase delay functions for the different values of θ can be seen in Fig. 1c and 1d. The behaviour is analogous to that of the nonseasonal model.

It is well known that the Wiener–Kolmogorov filters of the TSA corresponding to an infinite realization $\ldots, z_{-1}, z_0, z_1, \ldots$, are of the form $H(F)H(B)$, where B and F are the backshift and forward operators and $H(B)$ is a quotient of polynomials in B. To facilitate the comparison of the BND with the TSA, in Fig. 2a and 2b the gain functions of the $H(B)$ filters obtained with the TSA are presented for the same values of θ than in Fig. 1c and 1d. No phase delay functions are presented because in this case they are zero, given that the filter $H(F)H(B)$ is symmetric. Note that in all cases the gain functions are always less than or equal to one and that they are more in agreement with a gain function of a low–pass filter.

Finally, it has to be emphasized that the filters we have considered in this section correspond to an infinite realization, in the case of the TSA, and to a semi–infinite sample in the case of the BND. Since in practice only a finite series is available, the finite versions of these filters will differ from the infinite ones, especially at both ends of the series. This is an interesting topic which has been touched upon by, for example, Baxter and King (1995), but which we have not pursued in this paper for lack of space.

5 Application to the German Unemployment Series

To compare the properties of the TSA with those of the proposed BND, the German unemployment series running from 1962(I) to 1988(IV) is considered. This series has also been analyzed in Breitung (1994) and is selected to illustrate the main features of the two approaches. Needless to say that by considering a single example it is not possible to draw ultimate conclusions with respect to the merits or demerits of both approaches.

Applying the automatic model identification procedure of the software package TRAMO to the original time series[2], the following model was selected

[2]The series has not been logarithmically transformed or adjusted for outliers to facilitate the comparison with the BND

and estimated

$$(1 - 0.523B)(1 - B)(1 - B^4)z_t = (1 - 0.385B^4)a_t. \qquad (14)$$

The automatic model identification method of TRAMO proceeds in two stages. First, using an autoregressive model $\mathrm{AR}(2)(1)_n$ and $\mathrm{ARMA}(1,1)(1,1)_n$ models, where n is the number of seasons in the year, the unit roots are estimated. Then, the BIC criterion is applied to the differenced series to select among a wide range of multiplicative seasonal models a suitable one.

The fit is acceptable, although the residuals show some departure from normality due probably to the presence of some outlier. The components estimated with TSA are depicted in Fig. 3 and those estimated for the proposed BND with the two algorithms proposed in this paper (the results obtained with both algorithms are practically identical) in Fig. 3 b), d), f). Both estimated trends are quite similar, although the trend estimated with the BND is not so smooth. This is no surprise since the filters used by SEATS are two–sided whereas the ones used by the BND are one–sided. Also, the autoregressive factor $1 - 0.523B$ is assigned to the trend component in the TSA, whereas it is assigned to the stationary component in the BND.

It turns out that the recession of 1967 led to a sharp increase of cyclical (short–term) unemployment, while both the recession following the oil shock of 1973 and the recession of 1982–1984 caused a sharp raise in cyclical *and* long–term unemployment.

The seasonal components estimated with the BND and the TSA are very similar. However, the irregular component from TSA and the stationary component of the BND look quite different. This is explained, among other things, by the fact that, as mentioned earlier, the autoregressive factor $1 - 0.523B$ is assigned to the trend in the TSA and to the stationary component in the BND.

6 Conclusions

In this paper, the decomposition originally proposed by Beveridge and Nelson (1981) has been extended to arbitrary multiplicative seasonal ARIMA models. The proposed decomposition is based on a partial fraction expansion of the model followed by the series, where the regular and the seasonal unit roots of the differencing operator are assigned, respectively, to the trend and the seasonal component. The stationary autoregressive roots are assigned to a so–called stationary component, which may exhibit cyclical behaviour.

Two equivalent algorithms are proposed to compute the estimates of the components in the generalized BND. The algorithms are optimal in the mean squared sense and one of them, the augmented Kalman filter plus augmented smoothing, can also give the standard errors. However, the other algorithm is much simpler to apply.

The filters obtained with the BND seem to have less desirable features than the ones given by the TSA. First, their gain functions often take values greater than one and, second, they present a non–negligible phase delay effect. For the airline model with moving average factors of the form $1+\theta B$, where θ is positive, the trend filters given by the BND are unusable because the gain function takes values much greater than one.

Two possible advantages of the BND are that the decomposition always exists and that there are no revisions. However, this second advantage is more apparent than real because the lack of revisions comes at the expense of an increase in the error with which the components are estimated.

Appendix

In the expression of section 3 for x_1, δ models uncertainty with respect to the initial conditions and its distribution is unknown. Therefore, the ordinary Kalman filter cannot be applied and some device has to be used to handle δ, which can be considered as a vector of nuisance random variables. Kalbfleisch and Sprott (1970) proposed several methods to eliminate the dependence of the likelihood on nuisance parameters, which are also valid in the present context. More specifically, the marginal likelihood, which is the likelihood of a transformation of the data to eliminate the nuisance parameters, is the approach proposed by Ansley and Kohn (1985). The Bayesian approach, which consists of considering δ diffuse, is the approach of De Jong (1991).

For algorithmical purposes, we will use the approach of De Jong (1991) in this paper. Using the transition equation (13), we have the following lemma, whose proof is straightforward and is omitted.

Lemma 2 *Suppose that the series $z = (z_1, \ldots, z_N)'$ has been generated by the state space model* (12) *and* (13), *where $x_1 = A\delta + \Xi U$, as described earlier, and assume that δ is independent of the innovations a_t. Then, the following representation holds*

$$z = X\delta + \epsilon, \tag{15}$$

where, partitioning $X = (X_1, \ldots, X_N)'$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_N)'$ conforming to $z = (z_1, \ldots, z_N)'$, the X_t' and ϵ_t, $t = 1, \ldots, N$, can be obtained from the recursions $X_t' = H'J_t$ and $J_{t+1} = FJ_t$, with the initial condition $J_1 = A$, and $\epsilon_t = H'\eta_t$ and $\eta_{t+1} = F\eta_t + Ga_{t+1}$, with the initial condition $\eta_1 = \Xi U$. Besides, $E(\epsilon) = 0$, and $Cov(\delta, \epsilon) = 0$.

Let $\mathrm{Var}(a_t) = \sigma_a^2$ and $\mathrm{Var}(\epsilon) = \sigma_a^2\Sigma$ in (15). Following De Jong (1991), suppose that δ is independent of the $\{a_t\}$, has mean 0 and covariance matrix $\sigma_a^2 C$, and take the limit $C^{-1} \to 0$ to make it diffuse. Assuming normality in a_t and δ and letting $l(z)$ be the log–likelihood of z in (15) it is shown in De Jong (1991) that, apart from a constant, as $C^{-1} \to 0$,

$$\begin{aligned} l(z) + \frac{1}{2}\ln|\sigma_a^2 C| \to & -\frac{1}{2}\{(N - d^*)\ln(\sigma_a^2) + \ln|\Sigma| + \ln|X'\Sigma^{-1}X| \\ & + (z - X\hat{\delta})'\Sigma^{-1}(z - X\hat{\delta})/\sigma_a^2\}, \end{aligned} \tag{16}$$

where $\hat{\delta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}z$ and the mean squared error (Mse) of $\hat{\delta}$ is $\mathrm{Mse}(\hat{\delta}) = \sigma_a^2(X'\Sigma^{-1}X)^{-1}$. The limit expression in (16) is the diffuse log–likelihood. The parameter σ_a^2 can be concentrated out of the diffuse log–likelihood by replacing σ_a^2 in (16) with its maximum likelihood estimator $\hat{\sigma}_a^2 = (z - X\hat{\delta})'\Sigma^{-1}(z - X\hat{\delta})/(N - d^*)$.

The previous result tells us that making δ diffuse implies that (15) can be considered as a generalized linear regression model (GLS), where δ is the vector of regression parameters and $\hat{\delta}$ and $\hat{\sigma}_a^2$ are the GLS estimators.

In order to evaluate the diffuse log–likelihood efficiently, let $\Sigma = LL'$, with L lower triangular, be the Cholesky decomposition of $\Sigma = \mathrm{Var}(\epsilon)/\sigma_a^2$ and suppose that an efficient algorithm exists to compute $L^{-1}z$, $L^{-1}X$ and $|L|$. This algorithm is a slight modification of the DKF, which will be described later. Then, premultiplying (15) by L^{-1}, it is obtained that

$$L^{-1}z = L^{-1}X\delta + L^{-1}\epsilon, \tag{17}$$

where $\mathrm{Var}(L^{-1}\epsilon) = \sigma_a^2 I_N$. Therefore, model (17) is an ordinary linear regression model. The GLS estimators $\hat{\delta}$ and $\hat{\sigma}_a^2$ can now be efficiently and accurately obtained using the QR algorithm, as suggested by Kohn and Ansley (1985). This last algorithm premultiplies both $L^{-1}z$ and $L^{-1}X$ by an orthogonal matrix Q to obtain $v = QL^{-1}z$ and $(R', 0')' = QL^{-1}X$, where R is a nonsingular $d^* \times d^*$ upper triangular matrix. Then, $\hat{\delta} = R^{-1}v_1$ and $\hat{\sigma}_a^2 = v_2'v_2/(N - d^*)$, where $v = (v_1', v_2')'$, v_1 has dimension d^* and v_2 has dimension $N - d^*$. $|X'\Sigma^{-1}X|$ in (16) can be calculated as $|X'\Sigma^{-1}X| = |R'R|$.

To understand the meaning of the DKF of De Jong (1991), consider first that δ is zero in (15) and $\sigma_a^2 = 1$. Then, we can apply the ordinary Kalman filter, given by the recursions

$$\begin{aligned} e_t &= z_t - H'\hat{x}_{t|t-1}, \qquad \sigma_{t|t-1}^2 = H'\Sigma_{t|t-1}H \\ K_t &= F\Sigma_{t|t-1}H/\sigma_{t|t-1}^2, \quad \hat{x}_{t+1|t} = F\hat{x}_{t|t-1} + K_t e_t \\ \Sigma_{t+1|t} &= (F - K_t H')\Sigma_{t|t-1}F' + GG', \end{aligned}$$

where the initial conditions are $\hat{x}_{1|0} = 0$ and $\Sigma_{1|0} = \Xi\mathrm{Var}(U)\Xi'$ and the covariance matrix $\mathrm{Var}(U)$ can be efficiently computed using the results in Jones (1980). The sequence of standardized innovations $e_t/\sigma_{t|t-1}$, $t = 1, \ldots, N$ is an orthogonal sequence with mean zero and covariance matrix equal to the identity matrix. This implies that this sequence coincides with $L^{-1}z$ in (17). Also, $|L| = \prod_{t=1}^{N} \sigma_{t|t-1}$. These are standard results of the Kalman filter. Proofs can be seen in Anderson and Moore (1979).

A consequence of these results is that the Kalman filter can be seen as an algorithm that, applied to any vector v of data, yields $L^{-1}v$. Therefore, if δ is not zero in the GLS model (15), we can apply the Kalman filter to the data z and the columns of the X matrix to obtain $L^{-1}z$ and $L^{-1}X$. The DKF is an algorithm that allows for the automatic computation of these quantities. In this algorithm, the recursions for e_t and $\hat{x}_{t|t-1}$ in the Kalman filter are augmented to matrix recursions

$$\begin{aligned} (e_t, E_t) &= (z_t, 0) - H'(\hat{x}_{t|t-1}, \hat{X}_{t|t-1}), \\ (\hat{x}_{t+1|t}, \hat{X}_{t+1|t}) &= F(\hat{x}_{t|t-1}, \hat{X}_{t|t-1}) + K_t(e_t, E_t), \end{aligned}$$

where the additional columns correspond to new states for the columns of the X matrix. The other recursions in the Kalman filter remain the same and the initialization is $(\hat{x}_{1|0}, \hat{X}_{1|0}) = (0, -A)$ and $\Sigma_{1|0}$ as before. It can be shown, using the results in De Jong (1991), that stacking the vectors $(e_t, E_t)/\sigma_{t|t-1}$ one on top of the other for $t = 1, \ldots, N$, the matrix $(L^{-1}z, L^{-1}X)$ is generated.

The DKF also has the recursion $Q_{t+1} = Q_t + (e_t, E_t)'(e_t, E_t)/\sigma^2_{t|t-1}$, initialized with $Q_1 = 0$. This recursion accumulates the partial squares and cross products in such a way that

$$Q_{N+1} = \begin{bmatrix} (L^{-1}z)' \\ (L^{-1}X)' \end{bmatrix} \begin{bmatrix} L^{-1}z, L^{-1}X \end{bmatrix} = \begin{bmatrix} z'\Sigma^{-1}z & z'\Sigma^{-1}X \\ X'\Sigma^{-1}z & X'\Sigma^{-1}X \end{bmatrix}$$

and from Q_{N+1} the GLS estimators $\hat{\delta}$ and $\hat{\sigma}^2_a$ can be computed. We propose in this paper a Kalman filter algorithm which is the DKF without the recursion for Q_t and which applies instead the QR algorithm to $(L^{-1}z, L^{-1}X)$, in the manner described above. We think that this procedure is numerically more stable than solving the normal equations to obtain the GLS estimators and is not computationally expensive.

Note that σ^2_a is supposed to be one in the proposed algorithm because it can be estimated later with GLS.

Once the GLS estimators $\hat{\delta}$ and $\hat{\sigma}^2_a$ in (15) have been obtained, it can be shown, using the results in De Jong (1991), that the diffuse predictors $\hat{z}_{N+1}$ and $\hat{x}_{N+1}$ of z_{N+1} and X_{N+1} are

$$\hat{z}_{N+1} = H'\hat{x}_{N+1|N} + E_{N+1}\hat{\delta}, \quad \hat{x}_{N+1} = \hat{x}_{N+1|N} - \hat{X}_{N+1|N}\hat{\delta}$$
$$Mse(\hat{z}_{N+1}) = \hat{\sigma}^2_a \sigma^2_{t|t-1} + E_{N+1} Mse(\hat{\delta}) E'_{N+1}$$
$$Mse(\hat{x}_{N+1}) = \hat{\sigma}^2_a \Sigma_{N+1|N} + \hat{X}_{N+1|N} Mse(\hat{\delta}) \hat{X}'_{N+1|N},$$

where $\text{Mse}(\hat{\delta}) = \hat{\sigma}^2_a (X'\Sigma^{-1}X)^{-1} = \hat{\sigma}^2_a R^{-1} R'^{-1}$.

Diffuse smoothing refers to the process of obtaining the estimator $\hat{x}_t$ of the state x_t based on the entire data vector $z = (z_1, \ldots, z_N)'$. The estimator $\hat{x}_t$ can be obtained by means of an augmented version of any of the existing algorithms for smoothing, like the fixed point smoother or the fixed interval smoother. In this paper we will use an augmented fixed point smoother because it can be simplified so that very small storage requirements are needed, see Gómez and Maravall (1994), and because it is well suited for revisions of the estimates as new data come in.

The augmented fixed point smoother for x_s, $1 \le s \le N$, is the set of recursions

$$K^a_t = \Sigma^a_{t|t-1} H / \sigma^2_{t|t-1}, \quad \Sigma^a_{t+1|t} = \Sigma^a_{t|t-1}(F - K_t H')'$$

$$(\hat{x}_{s|t}, \hat{X}_{s|t}) = (\hat{x}_{s|t-1}, \hat{X}_{s|t-1}) + K_t^a(e_t, E_t)$$
$$\Sigma_{s|t} = \Sigma_{s|t-1} - \Sigma_{t|t-1}^a H(K_t^a)',$$

initialized with $\Sigma_{s|s-1}^a = \Sigma_{s|s-1}$, where $\sigma_{t|t-1}^2$, K_t, (e_t, E_t), $(\hat{x}_{s|s-1}, \hat{X}_{s|s-1})$ and $\Sigma_{s|s-1}$ are produced by the proposed Kalman filter algorithm. It can be shown that the estimator $\hat{x}_s$ and its Mse are obtained from

$$\hat{x}_s = \hat{x}_{s|N} - \hat{X}_{s|N}\hat{\delta}, \quad Mse(\hat{x}_s) = \hat{\sigma}_a^2 \Sigma_{s|N} + \hat{X}_{s|N} Mse(\hat{\delta}) \hat{X}_{s|N}'.$$

References

[1] Akaike, H., (1974), "Markovian Representation of Stochastic Processes and its Applications to the Analysis of Autoregressive Moving Average Processes", *Annals of the Institute of Statistical Mathematics*, **26**, 363–387.

[2] Anderson, B.D.O., and Moore, J.B., (1979), *Optimal Filtering*, Englewood Cliffs: Prentice Hall.

[3] Ansley, C.F. and Kohn, R., (1985) "Estimation, Filtering and Smoothing in State Space Models with Incompletely Specified Initial Conditions", *Annals of Statistics*, **13**, 1286–1316.

[4] Baxter, M. and King, R.G., (1995) "Measuring Business Cycles: Approximate Band–pass Filters for Economic Time Series", National Bureau of Economic Research, working paper N. 5022.

[5] Bell, W., (1984), "Signal Extraction for Nonstationary Series", *The Annals of Statistics*, **12**, 646-664.

[6] Bell, W., and Hillmer, S., (1991), "Initializing the Kalman Filter for Nonstationary Time Series Models", *Journal of Time Series Analysis*, **12**, 283–300.

[7] Beveridge, S. and Nelson, C.R., (1981), "A new Approach to Decomposition of Economic Time Series Into Permanent and Transitory Components With Particular Attention to Measurement of the Business Cycle", *Journal of Monetary Economics*, **7**, 151–174.

[8] Box, G.E.P., Hillmer, S.C., and Tiao, G.C., (1978), "Analysis and Modeling of Seasonal Time Series", in Zellner, A. (ed.), *Seasonal Analysis of Economic Time Series*, Washington, D.C.: U.S. Dept. of Commerce–Bureau of the Census, 309–334.

[9] Box, G.E.P., and Jenkins, G.M., (1976), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

[10] Breitung, J., (1994), A Model Based Seasonal Adjustment Method Using the Beveridge–Nelson Decomposition", *Allgemeines Statistisches Archiv*, **78**, 365–385.

[11] Burman, J.P., (1980), "Seasonal Adjustment by Signal Extraction", *Journal of the Royal Statistical Society* A, **143**, 321–337.

[12] Cuddington, J.T. and Winters, L.A., (1987), "The Beveridge–Nelson Decomposition of Economic Time Series: A Quick Computational Method", *Journal of Monetary Economics*, **19**, 125–127.

[13] De Jong, P., (1991), "The Diffuse Kalman Filter", *The Annals of Statistics*, **19**, 1073-1083.

[14] Gómez, V., and Maravall, A., (1994), "Estimation, Prediction and Interpolation for Nonstationary Series With the Kalman Filter", *Journal of the American Statistical Association*, **89**, 611-624.

[15] Gmez, V., and Maravall, A., (1997), "Programs TRAMO and SEATS", Instructions for the User (Beta Version: June 1997), Working Paper N. 97001, Direccin General de Anlisis y P.P., Ministry of Economy and Finance.

[16] Harvey, A.C. and Todd, P.H.J., (1983), "Forecasting Economic Time Series With Structural and Box–Jenkins Models: A Case Study", *Journal of Business and Economic Statistics*, **1**, 299–315.

[17] Jones, R., (1980), "Maximum Likelihood Fitting of ARMA Models to Time Series With Missing Observations", *Technometrics*, **22**, 389-395.

[18] Kalbfleisch, J. D., and Sprott, D. A., (1970), "Application of Likelihood Methods to Models Involving Large Numbers of Parameters" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, **32**, 175-208.

[19] Kohn, R., and Ansley, C.F., (1985), "Efficient Estimation and Prediction in Time Series Regression Models", *Biometrika*, **72**, 694–697.

[20] Maravall, A., and Pierce, D.A., (1987), "A Prototypical Seasonal Adjustment Model", *Journal of Time Series Analysis*, **8**, 177–193.

[21] Miller, S.M., (1987), "The Beveridge–Nelson Decomposition of Economic Time Series: Another Economical Computational Method", *Journal of Monetary Economics*, **19**, 125–127.

[22] Nerlove, M., Grether, D., and Carvalho, J.L., (1979), *Analysis of Economic Time Series: a Synthesis*, Academic Press, New York.

[23] Newbold, P., (1990), "Precise and Efficient Computation of the Beveridge–Nelson Decomposition of Economic Time Series", *Journal of Monetary Economomics*, **26**, 453–457.

[24] Newbold, P., and Vougas, D., (1995), "Beveridge–Nelson Type Trends for I(2) and Some Seasonal Models", *Journal of Time Series Analysis*, **17**, 151–170.

[25] Van Der Waerden, B.L., (1970), *Algebra* (translation of *Moderne Algebra*), Ungar, New York.

Fig. 1: a) Gain and b) Phase Delay Functions of Trend Filters (Nonseasonal BND). c) Gain and d) Phase Delay Functions of Trend Filters (Seasonal BND).

Fig. 2: a) Gain Functions of Trend Filters (Nonseasonal TSA).
. b) Gain Functions of Trend Filters (Seasonal TSA).

Fig. 3: a) Trend, c) Seasonal and e) Irregular component (TSA).
b) Trend, d) Seasonal and f) Stationary component (BND).