Working Paper 98-63
Statistics and Econometrics Series 30
September 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax 34 - 91- 624.9849

A BEVERIDGE-NELSON DECOMPOSITION FOR FRACTIONALLY INTEGRATED TIME SERIES

Miguel A. Ariño and Francesc Marmol*

Abstract

The purpose of this paper is to present a decomposition into trend or permanent component and cycle or transitory component of a time series that follows a nonstationary autoregressive fractionally integrated moving average *(ARFIMA(p,d,q))* model. As a particular case, for $d=1$ we obtain the well known Beveridge-Nelson decomposition of a series. For $d=2$ we get the decomposition of an I(2) series given by Newbold and Vougas (1996). The decomposition depends only on past data and is thus computable in real time. Computational issues are also discussed.

Key Words

Beveridge-Nelson decomposition; ARFIMA processes; Computation.

*I.E.S.E., Barcelona; Department of Statistics y Econometrics, Universidad Carlos III de Madrid, e-mail: fmarmol@est-econ.uc3m.es. The first author thanks Centro International de Investigaciones Financieras (CIIF) for its financial support.

1 Introduction

It is a well-known fact that many economic time series are nonstationary, but present an upward trend. During a long time this trend has been modelled by means of a simple linear time trend of the form

(1) $$y_t = \alpha + \beta t + s_t$$

where y_t usually stands for the natural log of the original series under study and the deviation from the linear trend, s_t, denotes a stationary process. Such an assumption is controversial, however, since it implies deterministic long-run growth rates.

In their seminal paper, Nelson and Plosser (1982) presented evidence that model (1) is inadequate for explaining the evolution of many economic time series, and proposed the alternative model

(2) $$y_t = \beta + y_{t-1} + s_t .$$

This model is known as integrated of order one, $I(1)$, process. It implies that the level of the series has a unit root whereas that the first differences of the series is stationary. Correspondingly, a process is $I(d)$ if we achieve stationarity after d integer differences.

When dealing with integrated processes, the underlying constituent components of the series, typically designated in the corresponding literature as "trend", "seasonal" and "irregular", are stochastic. In this case, Newbold (1991) showed that there is not a unique possible definition of "local trend". One possible characterization of trend is the well-known *Beveridge-Nelson (BN) decomposition.*

For time series following any (non-seasonal) $ARIMA(p,1,q)$ process, Beveridge and Nelson (1981) found a way of decomposing it into a trend (or permanent component) and a cycle (or transitory component). The permanent component or long-term trend of the series is simply defined as the value the series would have if it were on its long-run path in the current time period, and it is also the current observed value plus all forecastable future changes in the series beyond its mean rate of drift, so that, at any point in time, such a trend depends only on the information available to agents at that time. The permanent component of the *BN* decomposition of an $I(1)$ series turns out to be an $I(1)$ series while the cyclical component is a stationary series, as proved in Beveridge and Nelson (1981) or in Newbold and Vougas (1996).

Recently, Newbold and Vougas (1996) extend the *BN* decomposition to $I(2)$ series. In this case, the permanent component turned out to be an $I(2)$ series, so that, in contrast with the previous case, the slope of the permanent component evolves itself over time. The cyclical part of the series, in turn, remains stationary.

On the other hand, it has been argued many times (see, e.g., Diebold and Rudebusch, 1989) that a series may not follow neither model (1) nor be well represented as an integrated process. In this sense, the spectrum of many economic time series seems to be infinite at the origin, which suggests that some difference should be taken. However, after differencing, such time series appear not to have power at the origin. See, e.g., Granger (1969). In those cases, Granger and Joyeux (1980) and Hosking (1981) proposed the use of the so-called autoregressive fractionally integrated moving average (*ARFIMA*) models, a generalization of the standard *ARIMA* processes, where now the number d of differences is allowed to take any real value.

Fractionally integrated processes have received an increasing attention because of their ability to provide a natural and flexible characterization of the nonstationary and persistent characteristics of economic time series. They allow for more parsimonious models. Moreover, by allowing a rich range of spectral behaviour near the origin, they can provide superior approximations to the Wold representations of many economic time series. See Baillie (1996) for a review of the growing literature of econometric work on fractional processes and their applications in economics and finance.

The aim of this paper is to extend the *BN* decomposition to cover the fractional framework. From Baillie's (1996) survey, it seems quite reasonable to assume that most economic time series achieve stationarity after applying a fractional filter. Thus, it appears equally reasonable to look for a procedure describing the permanent and transitory components of the economic time series of interest under the fractional assumption. For this the paper is organised as follows. In Section 2 we extend the *BN* decomposition to nonstationary $ARIMA(p,d,q)$ models. As particular cases, when $d=1$ and $d=2$ we obtain the decompositions proposed by Beveridge and Nelson (1981) and Newbold and Vougas (1996), respectively. Section 3 considers the *BN* decomposition under a fractional set-up. We prove that such a decomposition exists and the components are functions of current and past, but not future, values of the series, so that they are computable in real time. Computational issues are gathered in Section 4 whereas Section 5 of the paper concludes. Proofs are given in the Appendix.

2 The *BN* decomposition for ARIMA(p,d,q) models

In their seminal paper, Beveridge and Nelson (1981) proved that if a time series, z_t, follows an $ARIMA(p,1,q)$ model, possibly with drift, then such a series can be decomposed into a trend and a cyclical components, denoted $\bar{z}_t$ and c_t respectively, and defined as

(3) $$\bar{z}_t = z_t + \sum_{j=1}^{\infty} \hat{x}_t(j)$$

and

(4) $$c_t = z_t - \bar{z}_t = -\sum_{j=1}^{\infty} \hat{x}_t(j),$$

with $w_t = z_t - z_{t-1}$, $x_t = w_t - \mu$, $\mu = E(w_t)$, $\phi(B)x_t = \theta(B)\varepsilon_t$, ε_t is white noise, B is the backshift operator, $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$, and with $\hat{x}_t(j)$ being the forecast at time t of x_{t+j} j periods ahead.

Therefore, the trend or permanent component $\bar{z}_t$ is defined as the current observed value of z_t plus all forecastable future changes in the series beyond the mean rate of drift μ. Moreover, it is quite simple to prove that the k-step ahead forecast of z_t at time t is

(5) $$\hat{z}_t(k) = z_t + \sum_{j=1}^{k} \hat{x}_t(j) + k\mu,$$

so that, asymptotically, as the forecast horizon k increases, this function $\hat{z}_t(k)$ is linear in k with slope μ and intercept

(6) $$z_t + \sum_{j=1}^{\infty} \hat{x}_t(j).$$

Hence, it is clear from expression (6) that the permanent component of a series as defined by Beveridge and Nelson (1981) is the value the series would have if it were on its long-run path in the current time period.

On the other hand, when z_t follows an $ARIMA(p,2,q)$ model, Newbold and Vougas (1996) have recently proved that the k-step ahead forecast of z_t made at time t is given by the expression

$$(7) \qquad \hat{z}_t(k) = (k+1)z_t - kz_{t-1} + \sum_{j=1}^{\infty}(k-j+1)\hat{x}_t(j),$$

where now $x_t = (1-B)^2 z_t$ and for simplicity and without loss of generality they assume that $\mu = 0$.

This function $\hat{z}_t(k)$, as the time horizon k increases, tends asymptotically to a linear function with slope

$$(8) \qquad \beta_t = z_t - z_{t-1} + \sum_{j=1}^{\infty}\hat{x}_t(j)$$

and intercept

$$(9) \qquad \alpha_t = z_t - \sum_{j=1}^{\infty}(j-1)\hat{x}_t(j).$$

Thus, as in the previous case, the forecasting function $\hat{z}_t(k)$ tends asymptotically to a linear function. However, now the slope of the asymptotic forecast function also evolves over time.

In this $ARIMA(p,2,q)$ case, Newbold and Vougas define the permanent part of the series again as the value that the series would have if it were on its long-run path in the current time period, that is, the permanent component is

(10) $$\bar{z}_t = z_t + \sum_{j=1}^{\infty}(1-j)\hat{x}_t(j),$$

and the cyclical component becomes

(11) $$c_t = z_t - \bar{z}_t = \sum_{j=1}^{\infty}(j-1)\hat{x}_t(j).$$

Following the approach of Newbold and Vougas, we extend the *BN* decomposition to series following (non-seasonal) $ARIMA(p,d,q)$ models for any positive integer *d*.

THEOREM 1. *Let* z_t *be a time series following an* $ARIMA(p,d,q)$ *model with positive integer d. Let* $x_t = (1-B)^d z_t$, *and let us assume for simplicity that* $\mu = E(x_t) = 0$. *Then*

(*i*) z_{t+k} *can be expressed as*

(12) $$z_{t+k} = P_{d-1}(k) + z_t + \sum_{i=1}^{k}\frac{(1-i)(2-i)\cdots(d-1-i)}{(d-1)!}x_{t+i},$$

where $P_{d-1}(k)$ *is a polynomial in k of degree* $d-1$, $P_{d-1}(k) = a_{d-1}k^{d-1} + \ldots + a_1 k + a_0$, *where the coefficients* a_i *are functions of* $z_t, z_{t-1}, \ldots, z_{t-d+1}, x_{t+1}, \ldots, x_{t+k}$.

(*ii*) *The forecasting function* $\hat{z}_t(k)$, *as the time horizon k increases, tends asymptotically to a* $d-1$ *degree polynomial in k with intercept equal to*

(13) $$z_t + \sum_{j=1}^{\infty}\frac{(1-j)(2-j)\cdots(d-1-j)}{(d-1)!}\hat{x}_t(j).$$

As a consequence of this result, we propose as a definition of the permanent component of the $ARIMA(p,d,q)$ process z_t the following expression

$$(14) \quad \bar{z}_t = z_t + \sum_{j=1}^{\infty} \frac{(1-j)(2-j)\cdots(d-1-j)}{(d-1)!} \hat{x}_t(j)$$

$$= z_t + (-1)^{d-1} \sum_{j=1}^{\infty} \frac{(j-1)(j-2)\cdots(j-d+1)}{(d-1)!} \hat{x}_t(j),$$

so that the cyclical component will be given by

$$(15) \quad c_t = z_t - \bar{z}_t = -\sum_{j=1}^{\infty} \frac{(1-j)(2-j)\cdots(d-1-j)}{(d-1)!} \hat{x}_t(j).$$

It is straightforward to show that these definitions give as particular cases the original *BN* decomposition for $ARIMA(p,1,q)$ models and the recent *BN* decomposition proposed by Newbold and Vougas (1996) for $ARIMA(p,2,q)$ data generating processes.

To close this section, it is worth mentioning that the interest of Theorem 1 does not rely in searching for a *BN* type decomposition for integrated processes of order greater than two, since one does not usually meet this kind of series in practice. The importance of the theorem stems in the fact that to understand how to find a *BN* decomposition of an $ARIMA(p,d,q)$ processes for any positive integer *d* does will provide with useful insights about how to proceed in order to obtain a *BN* decomposition in the $ARFIMA(p,d,q)$ case.

3 *ARFIMA* Models

When a given time series z_t becomes weakly stationary after differencing *d* times, and the degree of differentiation or *memory parameter*, *d*, is a real number, then the series is said to be *fractionally integrated of order d*, and it is said to follow an $ARFIMA(p,d,q)$ model if it can be represented as $\phi(B)(1-B)^d z_t = \theta(B)\varepsilon_t$, where $(1-B)^d$ is the fractionally difference operator defined as $(1-B)^d = 1+\sum_{j=0}^{\infty} d(d-1)(d-2)\cdots(d-j+1)(j!)^{-1}(-1)^j B^j$.

A fractionally integrated process is stationary if $d<\frac{1}{2}$ and all the roots of $\phi(z)=0$ lie outside the unit circle and and nonstationary if $d \geq \frac{1}{2}$, whereas it is invertible if $d>-\frac{1}{2}$ and all the roots of $\theta(z)=0$ lie outside the unit circle. It is both stationary and invertible if and only if $d \in \left(-\frac{1}{2},\frac{1}{2}\right)$ The problematic cases are when $d=\frac{1}{2}$ or $d=-\frac{1}{2}$, in which the series is either no stationary or no invertible.

In spite of being nonstationary, the process is mean-reverting with transitory memory, i.e., with any random shock having only a temporary influence on the series, if $d<1$, in contrast with the case when $d \geq 1$, where the process is both nonstationary and not mean-reverting with permanent memory, i.e., with any random shock having now a permanent effect on the present and future path of the series. On the other hand, a stationary fractionally integrated process has short-memory with autocorrelations decaying at an exponential rate if $d=0$, whereas it has long-memory with autocorrelations that die out at the slower hyperbolic rate if $0<d<1/2$.

A nonstationary fractionally integrated process can be made stationary by taking a suitable number of integer differences. The problematic cases are when $d=\frac{1}{2},\frac{3}{2},\frac{5}{2},\ldots$ In effect, for $d=n+\frac{1}{2}, d \geq \frac{1}{2}, n=1,2,3,\ldots$ after n differences in the series, an $ARFIMA\left(p,\frac{1}{2},q\right)$ series is obtained which is no stationary, and after $n+1$ differences the series obtained is $ARFIMA\left(p,-\frac{1}{2},q\right)$ which is not invertible.

In this section we are going to propose a decomposition into trend and cycle of a time series that follows a nonstationary $ARFIMA(p,d,q)$ model. Moreover, in order to rule out possible problems of lack of stationarity and/or invertibility, we shall be concerned here with values of the memory parameter d lying in the set $\partial = \{d \in \Re | d > \frac{1}{2}, d \neq n + \frac{1}{2}, n = 1, 2, 3, \ldots\}$. We focus the attention to the nonstationary case because it is the natural framework when dealing with *BN*-like-decompositions, but most of our results also hold in the stationary range.

Following the steps of Section 2, given a series z_t that follows an $ARFIMA(p,d,q)$ model with $d \in \partial$, we need to obtain an expression of z_{t+k} as a function of z_{t-j} for $j \geq 0$ and of x_{t+i} for $i = 1, \ldots, k$, where $x_t = (1-B)^d z_t$, $d \in \partial$. Afterwards, we will need to obtain the asymptotic behaviour of $\hat{z}_t(k)$ and proceed to separate the trend component and the cycle. The following theorem gives an expression of z_{t+k}.

<u>THEOREM 2</u>. *Let z_t be a time series following an $ARFIMA(p,d,q)$ model with $d \in \partial$. Let $x_t = (1-B)^d z_t$, and let us assume for simplicity that $\mu = E(x_t) = 0$. Then*

$$(16) \qquad z_{t+k} = \sum_{j=0}^{\infty} \phi_j(k) z_{t-j} + \psi(k),$$

where

$$(17) \qquad \phi_j(k) = \frac{\binom{k+j-1}{j}}{(k+j)!} d(1+d)(2+d)\cdots(k-1+d)(1-d)(2-d)\cdots(j-d)$$

$$(18) \qquad = \frac{\Gamma(k+d)\Gamma(j-d+1)}{j!(k-1)!(k+j)\Gamma(d)\Gamma(d+1)} = \frac{(-1)^j \Gamma(k+d)}{j!(k-1)!(k+j)\Gamma(d-j)}$$

for $d \neq n$, *and* $\psi(k)$ *is a function of* $x_{t+1}, x_{t+2}, \ldots + x_{t+k}$ *given by*

(19) $\psi(k) = \phi_0(k-1)x_{t+1} + \phi_0(k-2)x_{t+2} + \ldots + \phi_0(1)x_{t+k-1} + x_{t+k}$.

To decompose an *ARFIMA*(p,d,q) time series in permanent and cyclical component in the way of Beveridge and Nelson, the question turns out to be what is the function at which $\hat{z}_t(k)$ tends asymptotically. The value of this function at $k = 0$ would be the permanent component of the time series, and the remaining part the cyclical component. But this question has *no sense* in this context. In effect, when looking for an asymptote, one is customary looking for a straight line or a parabola or a curve within a known class of curves. In that sense, whatever real function f such that $lim_{k\to\infty}\left[\hat{z}_t(k) - f(k)\right] = 0$ is an asymptote of $\hat{z}_t(k)$. But for an asymptote to be of interest, it should be a function of a specific class of functions, and it seems there is no well-known function at which $\hat{z}_t(k)$ tends to.

Consequently, in order to propose a *BN*-like-decomposition of z_t in a fractional context, we observe that, according with Theorem 1, for any non-seasonal *ARIMA*(p,d,q) series with positive integer d the permanent component is

(20) $$\bar{z}_t = z_t + \sum_{j=1}^{\infty} f(d,j)\hat{x}_t(j),$$

where

(21) $$f(d,j) = \frac{1}{(d-1)!}(1-j)(2-j)\ldots(d-1-j) = \frac{(-1)^{d+1}(j-1)!}{(d-1)!(j-d)!}.$$

Values of $f(d,j)$ for different integer numbers d and j are given in Table 1. For a fix j the coefficients $f(d,j)$ are the coefficients of the terms x^k in the expression $(x-1)^{j-1}$ and thus Table 1 looks like the Tartaglia triangle.

f(d,j)	j						
d	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	0	-1	-2	-3	-4	-5	-6
3	0	0	1	3	6	10	15
4	0	0	0	-1	-4	-10	-20
5	0	0	0	0	1	5	15

Table 1. $f(d,j)$ for different values of d and j according to expression (21).

Our purpose in to extend Table 1 to values of $d \in \partial$. For this, our proposal for the permanent component $\bar{z}_t$ of an $ARFIMA(p,d,q)$ time series, would be the same expression (20) but with the $f(d,j)$ coefficients defined now for $d \in \partial$ in such a way that as d approaches to a positive integer n, $f(d,j)$ will approach to $f(n,j)$. Correspondingly, then, our definition of $f(d,j)$ for $d \in \partial$ is

$$(22)\quad f(d,j)=\begin{cases}\dfrac{(-1)^{d+1}(j-1)!}{(d-1)!(j-d)!} & \text{if } d=1,\,2,\,3,\cdots,\\[2ex] \dfrac{\Gamma(d-j)}{\Gamma(d)\Gamma(1-j+d-[d])} & \text{if } d\neq 1,\,2,\,3,\cdots, \text{ and } d\text{-}[d]<\tfrac{1}{2},\\[2ex] \dfrac{\Gamma(d-j)}{\Gamma(d)\Gamma(-j+d-[d])} & \text{if } d\neq 1,\,2,\,3,\cdots, \text{ and } d-[d]>\tfrac{1}{2},\end{cases}$$

with $[d]$ denoting the maximum integer number smaller or equal than d. $f(d,j)$ can also be defined as

$$(23)\quad f(d,j)=\begin{cases}\dfrac{(-1)^{d+1}(j-1)!}{(d-1)!(j-d)!} & \text{if } d=n=1,\,2,\,3,\cdots,\\[2ex] \dfrac{\Gamma(d-j)}{\Gamma(d)\Gamma(1-j+d-round(d))} & \text{otherwise,}\end{cases}$$

where round(d) stands for the integer number closest to d.

It is easy to check that with these definitions of $f(d,j)$, for each j

$$(24)\qquad \lim_{d\to n} f(d,j)=f(n,j)$$

as desired. Notice how $f(d,j)$ is not defined for $d=n+\frac{1}{2}$, and so, there is no *BN* decomposition for fractionally integrated processes with such a values of the memory parameter d. This is not a surprise, since for such cases there is no integer difference making the series both invertible and stationary. Nevertheless, the above comment should not represent a problem either, since the probability for the parameter d to be equal to $n+\frac{1}{2}$ for $d\in\partial$ is zero. Moreover, for each j and $d\in\partial$ the function $f(d,j)$ is continuous in $\Re\setminus\{n+\frac{1}{2}\}$. Table 2 shows values of $f(d,j)$ for different j and d according with expressions (22) or (23). Figures 2.*a*, 2.*b* and 2.*c*, in turn, plot $f(d,j)$ for different values of d while Figures 3.*a* and 3.*b* plot $f(d,2)$ and $f(d,3)$. It is clear from the graphics of Figures 2.*c*, 3.*a* and 3.*b* as well as from Table 2 that $f(d,j)$ is not continue for $d=n+\frac{1}{2}$.

f(d,j)	j						
d	1	2	3	4	5	6	7
0,6	0,672	0,672	0,672	0,672	0,672	0,672	0,672
0,9	0,936	0,936	0,936	0,936	0,936	0,936	0,936
1	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1,1	1,051	1,051	1,051	1,051	1,051	1,051	1,051
1,4	1,127	1,127	1,127	1,127	1,127	1,127	1,127
1,6	-0,448	-1,567	-2,686	-3,805	-4,924	-6,044	-7,163
1,9	-0,104	-1,144	-2,183	-3,223	-4,263	-5,303	-6,343
2	0,000	-1,000	-2,000	-3,000	-4,000	-5,000	-6,000
2,1	0,096	-0,860	-1,816	-2,771	-3,727	-4,682	-5,638
2,4	0,322	-0,483	-1,288	-2,093	-2,898	-3,703	-4,508
2,6	-0,168	0,392	2,350	5,708	10,464	16,620	24,174
2,9	-0,049	0,060	1,264	3,563	6,955	11,443	17,025
3	0,000	0,000	1,000	3,000	6,000	10,000	15,000

Table 2. $f(d,j)$ for different values of d and j according to expressions (22) or (23).

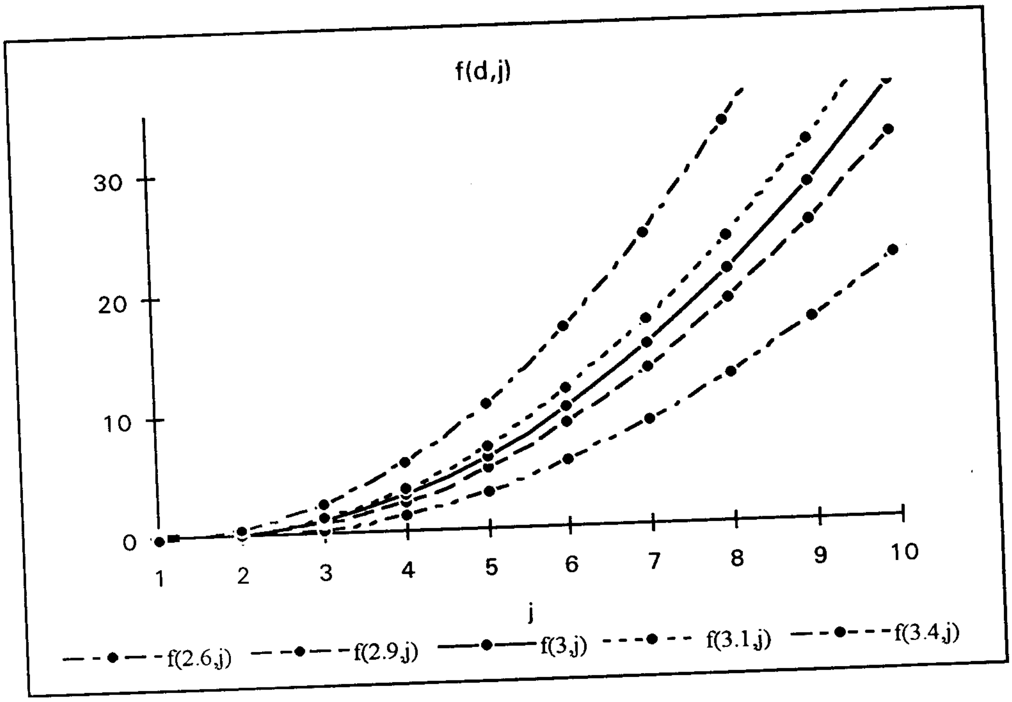

Figure 2.a. Graphic of $f(d,j)$ for $j = 2.6$, 2.9, 3, 3.1 and 3.4

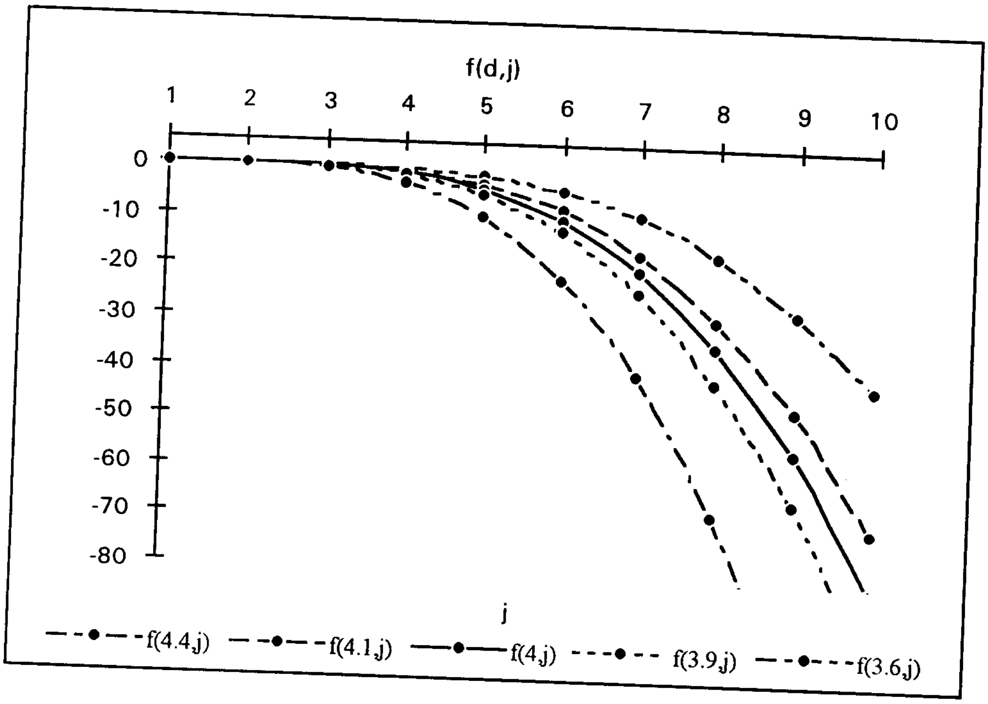

Figure 2.*b*. Graphic of $f(d,j)$ for $j = 3.6$, 3.9, 4, 4.1 and 4.4

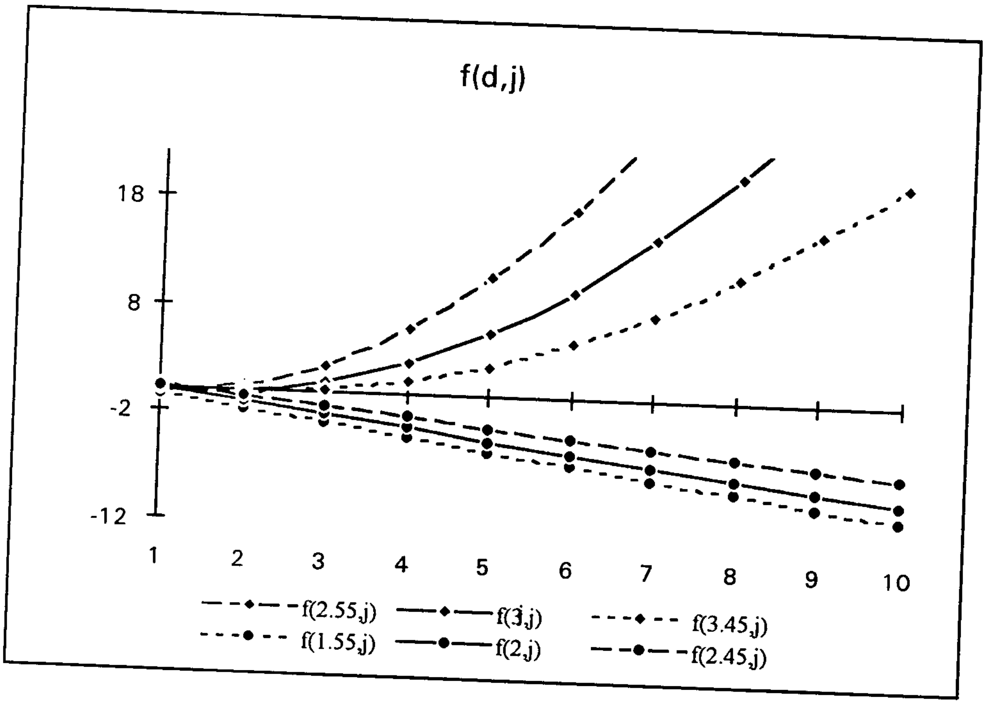

Figure 2.*c*. Graphic of $f(d,j)$ for $j = 1.55$, 2, 2.45, 2.55, 3 and 3.45

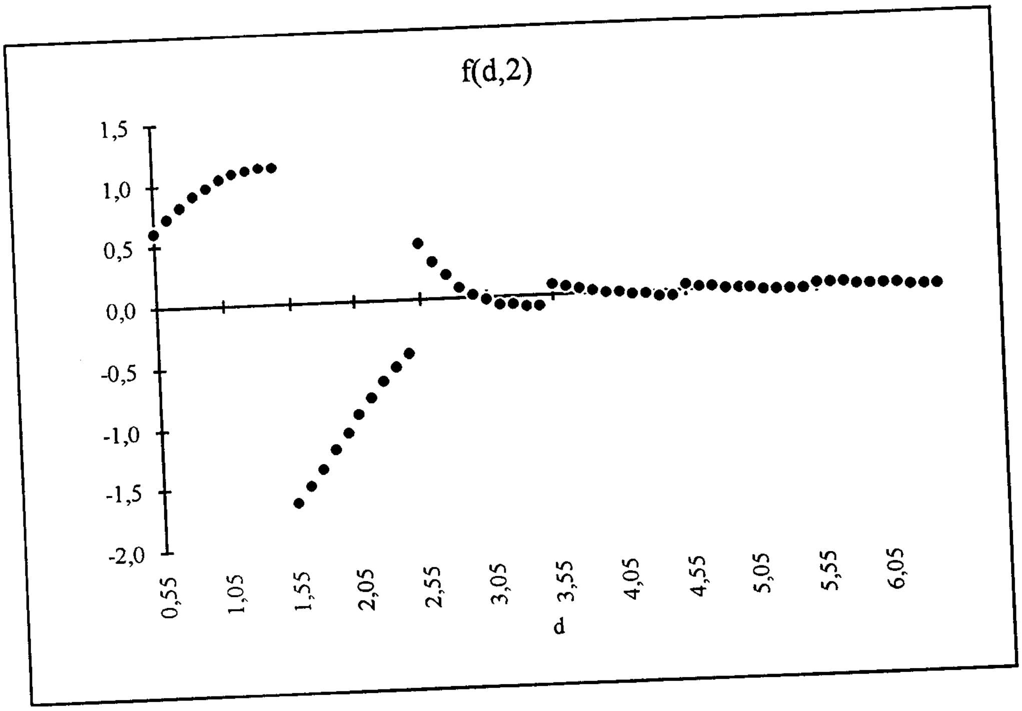

Figure 3.*a*. Values of $f(d,2)$.

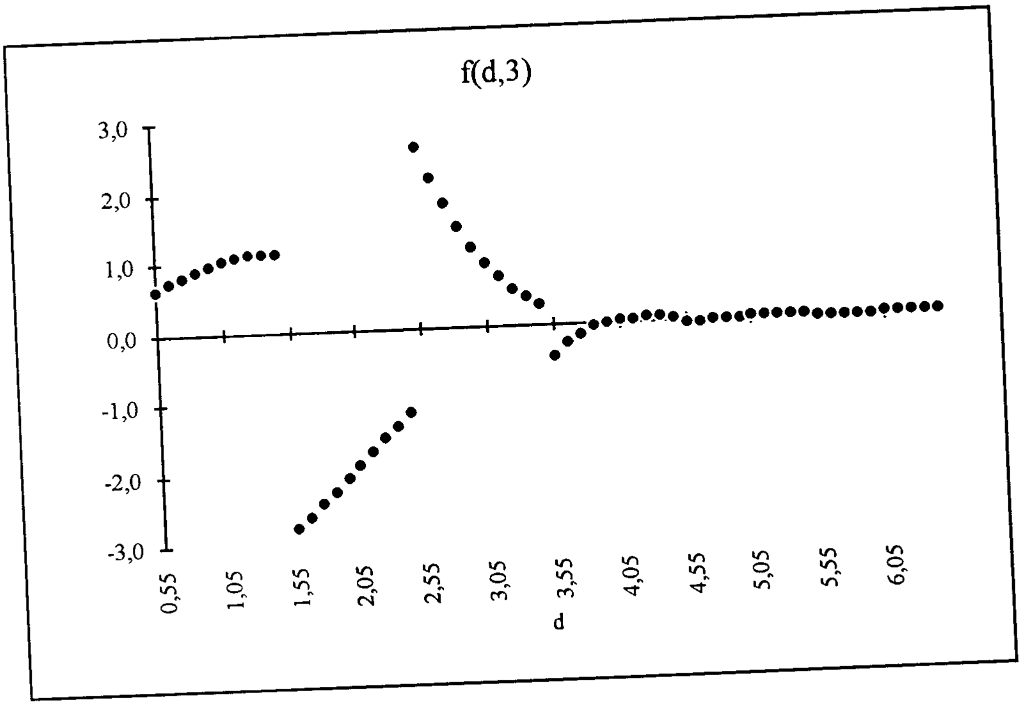

Figure 3.*b*. Values of $f(d,3)$.

4 Computational issues

Numerical calculations of the components in the *BN* decomposition of a fractionally integrated series require the computation of

(24) $$\sum_{j=1}^{\infty} f(d,j)\hat{x}_t(j).$$

To obtain an analytical expression of (24) we need the following result.

PROPOSITION 1. *Let $d \in \partial$ and k and q integer numbers with $k \geq 1$ and $q \geq 0$. Let A be a square matrix, and define $g_k(d,j) = \Gamma(d-j)/\Gamma(d-j-k+1)$ for $k \geq 2$ and $g_1(d,j) = 1$ for j positive, with $g_k(n,j) = \prod_{i=1}^{k-1}(n-j-i)$. Then,*

(25) $$S = S(d,k,q) = \sum_{h=1}^{\infty} g_k(d,q+h)A^h = \left[\sum_{i=1}^{k}(-1)^{i+1}\binom{k-1}{i-1}\frac{g_{k+1}(d,q)}{d-q-i}A^i\right](I-A)^{-k}.$$

We are now in position to obtain an expression for (24). Thus, as in Newbold (1990) and Newbold and Vougas (1996) consider the decomposition

(26) $$\sum_{j=1}^{\infty} f(d,j)\hat{x}_t(j) = \sum_{j=1}^{q} f(d,j)\hat{x}_t(j) + \sum_{j=q+1}^{\infty} f(d,j)\hat{x}_t(j).$$

Since for $j > q$ $\hat{x}_t(j) = \phi_1\hat{x}_t(j-1)+\ldots+\phi_p\hat{x}_t(j-p)$, we have that

(27) $$\hat{x}_t(q+h) = e'A^h x,$$

where e and x are $1 \times p$ vectors defined by $e' = (1,0,\ldots,0)$ and $x' = (\hat{x}_t(q),\ldots,\hat{x}_t(q-p+1))$ and $A = (a_{ij})$ is a $p \times p$ matrix with $a_{1j} = \phi_j$ for $1 \leq j \leq p$, $a_{i,i-1} = 1$ for $1 \leq i \leq p$ and all other elements are zero.

Now, for $\frac{1}{2} < d < \frac{3}{2}$, k is equal to 1 and $f(d,j) = 1/\Gamma(d)$, and expression (24) is calculated as in Newbold (1990). For $d > \frac{3}{2}$, k is greater than 1 and we can apply the previous proposition so that $f(d,j) = g_k(d,j)/\Gamma(d)$, and thus

$$\sum_{j=q+1}^{\infty} f(d,j)\hat{x}_t(j) = \frac{1}{\Gamma(d)}\sum_{h=1}^{\infty} g_k(d,q+h)\hat{x}_t(q+h)$$

$$= \frac{1}{\Gamma(d)}\sum_{h=1}^{\infty} g_k(d,q+h)\, e' A^h x = \frac{1}{\Gamma(d)} e'\left(\sum_{h=1}^{\infty} g_k(d,q+h) A^h\right) x$$

$$= \frac{1}{\Gamma(d)} e' S(d,k,q) x,$$

and therefore,

$$(28) \quad \sum_{j=1}^{\infty} f(d,j)\hat{x}_t(j) = \sum_{j=1}^{q} f(d,j)\hat{x}_t(j) + \frac{1}{\Gamma(d)} e' S(d,k,q) x .$$

When $d = 1$ or $d = 2$ and due to the fact that, for the special form of matrix A, $A(I - A)^{-1} = (I - A)^{-1} A$, expression (28) is the same as the obtained by Newbold (1990) and Newbold and Vougas (1996), respectively. When d is an integer number, then $k = d$ and we get $-C(t) = \sum_{j=1}^{\infty} f(d,j)\hat{x}_t(j) = \Gamma(d-q) e' A^k (I - A)^{-k} x/\Gamma(d)$. If, additionally $q = 0$, then $-C(t) = e' A^k (I - A)^{-k} x$.

5 Summary and Conclusions

In this article we have generalised the classical Beveridge-Nelson decomposition of an $ARIMA(p,1,q)$ and the recent $ARIMA(p,2,q)$ extension of Newbold and Vougas (1996) to nonstationary $ARFIMA(p,d,q)$ models, $d \in \partial$. This extension, by allowing for more flexible and natural characterizations of the trend-reversing behavior of a time series, can be proved useful when decomposing a time series into a permanent and a transitory components. Moreover, as in the *ARIMA* case, our decomposition depends only on past data and therefore is computable in real time.

To that purpose we have first identified how should be such a decomposition for $ARIMA(p,d,q)$ models for arbitrary integer d. Once it was understood how should be such a decomposition, we propose a decomposition for the more general *ARFIMA* framework with weights $f(d,j)$ in the permanent component selected via continuity arguments and in order to resemble the paths of the weights in the *ARIMA* case.

We have finished our paper by giving analytical formulas to numerically calculate both components of the fractionally integrated *BN* decomposition. Such a expressions allow efficient and exact computation of the *BN*-type trends discussed in this paper.

Appendix

PROOF OF THEOREM 1. For simplicity we shall first provide a formal proof for the particular $d = 4$ case. The steps to follow in order to prove the general $ARIMA(p,d,q)$ case for other values of *d* are completely similar to those of $d = 4$.

Therefore, assume that z_t is well represented as an $ARIMA(p,4,q)$ process and let $x_t = (1-B)^4 z_t$ so that

(A.1) $z_t = 4z_{t-1} - 6z_{t-2} + 4z_{t-3} - z_{t-4} + x_t$.

In this case, z_{t+k} can be expressed as

(A.2) $z_{t+k} = a(k)z_t + b(k)z_{t-1} + c(k)z_{t-2} + d(k)z_{t-3} + \psi(k)$,

where $\psi(k)$ is a function of $x_{t+k}, x_{t+k-1}, \dots, x_{t+1}$.

To see this it is enough to observe that

$$\begin{aligned}
z_{t+k} &= 4z_{t+k-1} - 6z_{t+k-2} + 4z_{t+k-3} - z_{t+k-4} + x_{t+k} \\
&= [4a(k-1) - 6a(k-2) + 4a(k-3) - a(k-4)]z_t \\
&+ [4b(k-1) - 6b(k-2) + 4b(k-3) - b(k-4)]z_{t-1} \\
&+ [4c(k-1) - 6c(k-2) + 4c(k-3) - c(k-4)]z_{t-2} \\
&+ [4d(k-1) - 6d(k-2) + 4d(k-3) - d(k-4)]z_{t-3} \\
&+ 4\psi(k-1) - 6\psi(k-2) + 4\psi(k-3) - \psi(k-4) + x_{t+k},
\end{aligned}$$

and thus,

$$a(k) = 4a(k-1) - 6a(k-2) + 4a(k-3) - a(k-4),$$

$$b(k) = 4b(k-1) - 6b(k-2) + 4b(k-3) - b(k-4),$$

$$c(k) = 4c(k-1) - 6c(k-2) + 4c(k-3) - c(k-4),$$

$$d(k) = 4d(k-1) - 6d(k-2) + 4d(k-3) - d(k-4),$$

$$\psi(k) = 4\psi(k-1) - 6\psi(k-2) + 4\psi(k-3) - \psi(k-4) + x_{t+k},$$

with initial conditions

$$a(-3) = 0, b(-3) = 0, c(-3) = 0, d(-3) = 1,$$

$$a(-2) = 0, b(-2) = 0, c(-2) = 1, d(-2) = 0,$$

$$a(-1) = 0, b(-1) = 1, c(-1) = 0, d(-1) = 0,$$

$$a(0) = 1, b(0) = 0, c(0) = 0, d(0) = 0,$$

$$\psi(-3) = \psi(-2) = \psi(-1) = \psi(0) = 0,$$

yielding

$$a(k) = \frac{1}{6}k^3 + k^2 + \frac{11}{6}k + 1,$$

$$b(k) = -\frac{1}{2}k^3 - \frac{5}{2}k^2 - 3k,$$

$$c(k) = -\frac{1}{2}k^3 - 2k^2 + \frac{3}{2}k,$$

$$d(k) = -\frac{1}{6}k^3 - \frac{1}{2}k^2 - \frac{1}{3}k,$$

and

$$\psi(k) = \sum_{i=1}^{k}\left[1 + \frac{1}{6}(k-i)^3 + (k-i)^2 + \frac{11}{6}(k-i)\right]x_{t+i}$$

$$= \frac{1}{6}\left(\sum_{i=1}^{k} x_{t+i}\right)k^3 + \left(\sum_{i=1}^{k}\left(1 - \frac{i}{2}\right)x_{t+i}\right)k^2 + \left(\sum_{i=1}^{k}\left(2i^2 - 2i + \frac{11}{6}\right)x_{t+i}\right)k$$

$$- \sum_{i=1}^{k} \frac{(i-1)(i-2)(i-3)}{6} x_{t+i} ,$$

so that

$$z_{t+k} = \left[\frac{1}{6} z_t - \frac{1}{2} z_{t-1} + \frac{1}{2} z_{t-2} - \frac{1}{6} z_{t-3} + \left(\sum_{i=1}^{k} x_{t+i} \right) \right] k^3$$

$$+ \left[z_t - \frac{5}{2} z_{t-1} + 2 z_{t-2} - \frac{1}{2} z_{t-3} + \left(\sum_{i=1}^{k} \left(1 - \frac{i}{2} \right) x_{t+i} \right) \right] k^2$$

$$+ \left[\frac{11}{6} z_t - 3 z_{t-1} + \frac{3}{2} z_{t-2} - \frac{1}{3} z_{t-3} + \left(\sum_{i=1}^{k} \left(2i^2 - 2i + \frac{11}{6} \right) x_{t+i} \right) \right] k$$

$$+ z_t - \sum_{i=1}^{k} \frac{(i-1)(i-2)(i-3)}{6} x_{t+i} ,$$

and so the first part of the theorem is proved.

To prove the second part, notice that $\hat{z}_t(k)$ tends asymptotically to a polynomial of degree 3 with intercept

$$\bar{z}_t = z_t - \sum_{j=1}^{\infty} \frac{(j-1)(j-2)(j-3)}{6} \hat{x}_t(j) .$$

Consequently, if the permanent component $\bar{z}_t$ of z_t is defined as the value the series would have if it were on its long-run path in the current time period, then this permanent component should be the intercept of this asymptotically polynomial, and so

$$\bar{z}_t = z_t + \sum_{j=1}^{\infty} \frac{(1-j)(2-j)(3-j)}{3!} \hat{x}_t(j) ,$$

and the theorem is proved for $d = 4$.

Finally, in order to prove the theorem for any integer d, we should follow similar steps as in the $d = 4$ case. So, if $x_t = (1-B)^d z_t$, then

$$z_t = \sum (-1)^{i+1} \binom{d}{i} z_{t-i} + x_t ,$$

and z_{t+k} can be expressed as

(A.3) $\quad z_{t+k} = a_0(k) z_t + a_1(k) z_{t-1} + \ldots + a_{d-1}(k) z_{t-d+1} + \psi(k)$.

In the same manner, it can be proved that now the $a_j(k)$ terms are polynomials of degree $d-1$ with $a_0(0) = 1$, and $a_i(0) = 0$ for $1 \le i \le d-1$, whereas $\psi(k)$ is a polynomial in k of degree $d-1$ whose coefficients are combinations of $x_{t+1}, \ldots x_{t+k}$, and with

$$\psi(0) = (-1)^{d-1} \sum_{i=1}^{k} \frac{(i-1)(i-2)\ldots(i-d+1)}{(d-1)!} x_{t+i} .$$

Rearranging terms in (A.3) we obtain that

$$z_{t+k} = P_{d-1}(k) + z_t + (-1)^{d-1} \sum_{i=1}^{k} \frac{(i-1)(1-2)\ldots(i-d+1)}{(d-1)!} x_{t+i}$$

with $P_{d-1}(0) = 0$, and so, the permanent component of the series in the general $ARIMA(p,d,q)$ case becomes

$$\bar{z}_t = z_t + (-1)^{d-1} \sum_{j=1}^{\infty} \frac{(j-1)(j-2)\ldots(j-d+1)}{(d-1)!} \hat{x}_t(j)$$

concluding the proof of the theorem. ■

PROOF OF THEOREM 2. Since $x_t = (1-B)^d z_t$ we have that

(A.4) $\quad z_t = \sum_{j=0}^{\infty} \alpha_j z_{t-j-1} + x_t$

being $\alpha_0 = d$ and $\alpha_j = d(1-d)(2-d)\cdots(j-d)/(j+1)!$. Assuming that for $s \le k$

$$z_{t+s} = \sum_{j=0}^{\infty} \phi_j(s) z_{t-j} + \psi(s) ,$$

we have that for $s = k + 1$

$$z_{t+k+1} = \sum_{j=0}^{\infty} \alpha_j z_{t+k-j} + x_{t+k+1} = \sum_{j=0}^{\infty} \alpha_j \left[\sum_{r=0}^{\infty} \phi_r (k-j) z_{t-r} + \psi(k-j) \right] + x_{t+k+1} =$$

$$\sum_{j=0}^{\infty} \left[\sum_{r=0}^{\infty} \alpha_r \phi_j (k-r) \right] z_{t-j} + \sum_{j=0}^{\infty} \alpha_j \psi(k-j) + x_{t+k+1} = \sum_{j=0}^{\infty} \phi_j (k+1) z_{t-j} + \psi(k+1),$$

and thus,

$$\phi_j (k+1) = \sum_{r=0}^{\infty} \alpha_r \phi_j (k-r),$$

and

$$\psi(k+1) = \sum_{j=0}^{\infty} \alpha_j \psi(k-j) + x_{t+k+1}.$$

The next step is to find the analytical expression of $\phi_j(k)$. Let us start with $j = 0$. We know that

$$\phi_0 (k+1) = \sum_{r=0}^{\infty} \alpha_r \phi_0 (k-r),$$

with initial conditions $\phi_0(0) = 1$, $\phi_0(m) = 0$ for $m < 0$. By recurrence it can prove that $\phi_0(1) = d$, $\phi_0(2) = (2!)^{-1} d(1+d)$,..., $\phi_0(k) = (k!)^{-1} d(1+d)(2+d)\cdots(k-1+d)$. The initial conditions for $j = 1$ are $\phi_1(0) = 0$, $\phi_1(-1) = 1$ and $\phi_1(m) = 0$ for $m < -1$. Thus, by recurrence it is obtained now that

$$\phi_1(k) = \frac{k}{(k+1)!} d(1-d)(1+d)(2+d)\cdots(k-1+d).$$

For general j the initial conditions are $\phi_j(-j) = 1$ and $\phi_j(m) = 0$ for $m \le 0$, $m \ne -j$, and the expression of $\phi_j(k)$ is

$$\phi_j(k) = \frac{\binom{k+j-1}{j}}{(k+j)!} d(1-d)(2-d)\cdots(j-d)(1+d)(2+d)\cdots(k-1+d).$$

This expression, in turn, for $d \in \partial$ and $d \neq n$ becomes equal to

$$\phi_j(k) = \frac{\Gamma(k+d)\Gamma(j-d+1)}{j!(k-1)!(k+j)\Gamma(d)\Gamma(d+1)} = \frac{(-1)^j\,\Gamma(k+d)}{j!(k-1)!(k+j)\Gamma(d-j)}.$$

Finally, to conclude the proof of the theorem only remains finding the explicit expression of $\psi(k)$. For this, notice from the fact that

$$\psi(k+1) = \sum_{j=0}^{\infty} \alpha_j \psi(k-j) + x_{t+k+1}$$

and the initial condition $\psi(m) = 0$ for $m \leq 0$ that

$$\psi(1) = x_{t+1},$$

$$\psi(2) = d x_{t+1} + x_{t+2} = \phi_0(1)x_{t+1} + x_{t+2},$$

$$\psi(3) = \frac{1}{2}d(d+1)x_{t+1} + d x_{t+2} + x_{t+3} = \phi_0(2)x_{t+1} + \phi_0(1)x_{t+2} + x_{t+3},$$

and in general

$$\psi(k) = \sum_{j=1}^{k} \phi_0(k-j)x_{t+j},$$

concluding the proof of the theorem. ■

PROOF OF PROPOSITION 1. In order to proof this proposition, we need some intermediate results collected in the following lemma.

<u>LEMMA 1</u>. *Under the same assumptions as in Proposition* 1,

(*i*) *for* $k > 1$, $g_k(d,q+h) - g_k(d,q+h-1) = (-k+1)g_{k-1}(d,q+h)$.

(*ii*) *For* $k > 1$, $g_k(d,q+1) + (k-1)g_{k-1}(d,q+1) = g_k(d,q)$.

(iii) For $k \geq 1$, $g_{k+1}(d,q) = g_k(d,q)(d-q+k)$,

and

(iv) for $1 \leq i \leq k$,

$$g_k(d,q)\binom{k-1}{i-1} + (-k+1)g_k(d,q)\binom{k-2}{i-2}\frac{1}{d-q-i} = g_{k+1}(d,q)\binom{k-1}{i-1}\frac{1}{d-q-i}.$$

PROOF OF LEMMA 1. Parts (*i*) to (*iii*) are straightforward using the definition of g_k. To prove (*iv*) note that, by using (*iii*) we have that

$$g_k(d,q)\binom{k-1}{i-1} + (-k+1)g_k(d,q)\binom{k-2}{i-1}\frac{1}{d-q-i}$$

$$= g_k(d,q)\binom{k-1}{i-1}\left[1 + \frac{(-k+1)(k-i)}{(d-q-i)(k-1)}\right] = g_{k+1}(d,q)\binom{k-1}{i-1}\frac{1}{d-q-i},$$

and part (*iv*) is proved.

Now, to prove the proposition, it is easy to check that expression (25) is true for $k = 2$. Assuming that it is also true for $k-1$, let us check that it also holds for *k*. For this, notice that, after some manipulations, we obtain

$$\text{(A.5)} \quad S - SA = g_k(d,q+1)A + \sum_{h=2}^{\infty}\left[g_k(d,k+h) - g_k(d,k+h-1)\right]A^h.$$

Using parts (*i*) and (*ii*) of the Lemma, this expression is equal to

$$g_k(d,q+1)A + (-k+1)\sum_{h=2}^{\infty} g_{k-1}(d,q+h)A^h$$

$$= g_k(d,q+1)A + (-k+1)\left[S(d,k-1,q) - g_{k-1}(d,q+1)A\right]$$

$$= \left[g_k(d,q+1) + (k-1)g_{k-1}(d,q+1)\right]A + (-k+1)S(d,k-1,q)$$

$$= g_k(d,q)A + (-k+1)S(d,k-1,q).$$

Since expression (25) is assumed to be true for $k-1$, this last expression becomes equal to

$$g_k(d,q)A + (-k+1)\left[\sum_{i=1}^{k-1}(-1)^i g_k(d,q)\binom{k-2}{i-1}\frac{1}{d-q-i}A^i\right](I-A)^{-k+1}$$

$$= g_k(d,q)A(I-A)^{k-1} + (-k+1)\left[\sum_{i=1}^{k-1}(-1)^i g_k(d,q)\binom{k-2}{i-1}\frac{1}{d-q-i}A^i\right](I-A)^{-k+1}$$

$$= \left[g_k(d,q)(-1)^{k+1}A^k + \sum_{i=1}^{k-1} g_k(d,q)(-1)^{i+1}\left[\binom{k-1}{i-1} + \binom{k-2}{i-1}\frac{-k+1}{d-q-i}\right]A^i\right](I-A)^{-k+1}.$$

By using parts (*iii*) and (*iv*) of Lemma 1, we can summarise the above expressions, yielding

$$\text{(A.6)} \quad S - SA = \left[\sum_{i=1}^{k}(-1)^{i+1}\binom{k-1}{i-1}\frac{g_{k+1}(d,q)}{d-q-i}A^i\right](I-A)^{-k+1},$$

and thus

$$\text{(A.7)} \quad S = \left[\sum_{i=1}^{k}(-1)^{i+1}\binom{k-1}{i-1}\frac{g_{k+1}(d,q)}{d-q-i}A^i\right](I-A)^{-k}. \blacksquare$$

References

Baillie, R. (1996), "Long memory processes and fractional integration in econometrics", *Journal of Econometrics* 73, 5-59.

Beveridge, S and C.R. Nelson (1981), "A new approach to the decomposition of economic time series into permanent and transient components with particular attention to measurement of the business cycle", *Journal of Monetary Economics* 7, 151-174.

Diebold, F.X. and G.D. Rudebusch (1989), "Long memory and persistence in aggregate output", *Journal of Monetary Economics* 24, 189-209.

Granger, C.W.J. (1969), "The typical spectral shape of an economic variable", *Econometrica* 34, 150-161.

Granger, C.W.J. and R. Joyeux (1980), "An introduction to long memory time series models and fractional differencing", *Journal of Time Series Analysis* 1, 15-39.

Hosking, J.R.M. (1981), "Fractional differencing", *Biometrika* 68, 165-176.

Nelson, C.R. and C. I. Plosser (1982), "Trends and random walks in macroeconomic time series: some evidence and implications", *Journal of Monetary Economics* 10, 139-162.

Newbold, P. (1990), "Precise and efficient computation of the Beveridge-Nelson decomposition of economic time series", *Journal of Monetary economics* 26, 453-457.

Newbold, P. (1991), "Structural decomposition of time series with implications in economics, accounting, and finance research", *Review of Quantitative Finance and Accounting* 1, 259-279.

Newbold, P. and D. Vougas (1996), "Beveridge-Nelson type trends for $I(2)$ and some seasonal models", *Journal of Time Series Analysis* 17, 151-169.