ISSN 1440-771X

MONASH UNIVERSITY
AUSTRALIA

DEPARTMENT OF ECONOMETRICS
AND BUSINESS STATISTICS

Persistence and Nonstationary Models

B. P. M. McCabe, G. M. Martin and A. R. Tremayne

Working Paper 16/2003

Persistence and Nonstationary Models*

B.P.M. McCabe†, G.M. Martin‡ and A.R. Tremayne§

September 19, 2003

Abstract

The aim of this paper is to examine the measurement of persistence in a range of time series models nested in the framework of Cramer (1961). This framework is a generalization of the Wold (1938) decomposition for stationary time series which, in addition to accommodating the standard I(0) and I(1) models, caters for alternative nonstationary processes. Three measures of persistence are considered, namely the long-run impulse response, variance ratio and autocorrelation functions. Particular emphasis is given to the behaviour of these measures in a range of nonstationary models. We document conflict that arises between different measures, applied to the same model, as well as conflict arising from the use of a given measure in different models. Precisely which persistence measures are time dependent and which are not, is highlighted. The nature of the general representation used also helps clarify what shock the impulse response function refers to in the case of models where more than one random disturbance impinges on the time series.

Keywords: Cramer Representation, Stochastic Unit Root Model, Stochastic Integration, Impulse Response, Variance Ratio, Autocorrelation Function, Long Memory.

JEL Classifications: C10, C22.

*This research has been supported by an Australian Research Council Discovery Grant, No. DP0208333. The authors would like to thank Don Poskitt for providing some useful comments.

†School of Management, University of Liverpool.

‡Department of Econometrics and Business Statistics, Monash University. Corresponding author: Department of Econometrics and Business Statistics, P.O. Box, 11E, Monash University, Victoria, 3800, Australia. *(Email: Gael.Martin@Buseco.monash.edu.au).*

§School of Economics and Political Science, University of Sydney.

1 Introduction

A time series is said to exhibit persistence if the impact of a shock to the series continues indefinitely into the future. Linked to this is the idea that a persistent series contains a permanent component, whereby past shocks exert an ongoing effect on the level of the series. In contrast, a series that is not persistent is said to contain a transitory component only, with the impact of shocks dissipating over time. The identification and measurement of persistence in an economic time series is crucial to an understanding of the response of the economic variable to shocks, in particular to policy-induced shocks. The presence of a large permanent component, for instance, implies that fiscal and monetary policies aimed at producing short-run corrections, may well have unintended long-run consequences. In fact, the relative magnitudes of the permanent and transitory components or, alternatively, the degree of persistence in the series, can be viewed as statistical evidence on the validity of competing business cycle theories, including the importance which they ascribe to interventionist government policies; see Campbell and Mankiw (1987, 1989) and Cochrane (1988) for seminal contributions.

Standard measures of persistence are: the long-run impulse response function; the long-run variance ratio function; and the long-run autocorrelation function. Whilst the impulse response function is a direct measure of the impact of a shock on the level of a time series, the other two measures provide a more indirect indication of persistence. The long-run variance ratio conventionally measures the proportion of variation in the first differenced time series that is attributable to the permanent component in the series, whilst the long-run autocorrelation function measures the correlation between the level of the time series in two periods that are infinitely far apart.

Traditionally, discussion of persistence in economic and financial time series has taken place in the context of models that are either integrated of order zero (I(0)) or of order one (I(1)). In this context, both the impulse response and variance ratio functions indicate, unambiguously, that I(1) models are persistent, whilst I(0) models are not. The lack of persistence of the I(0) process is further confirmed by the autocorrelation function, which decays as the time between observations increases. Extension of the impulse response measure to fractionally integrated (I(d)) models has lead to the conclusion that such models are nonpersistent if the fractional integration parameter, d, is less than unity and persistent otherwise. For stationary fractional models, with $d < 0.5$, a lack of persistence is indicated by the decay that occurs in both the impulse response and autocorrelation functions; see

Brockwell and Davis (1991) and Baillie (1996).

Linked to the concept of persistence is the concept of memory. Traditionally, memory is measured via the absolute sum of the autocorrelation coefficients, with a convergent sum indicating short (or intermediate) memory and a divergent sum indicating long memory (Brockwell and Davis, 1991, 13.2). As is clear, it is quite possible for a series to be nonpersistent, in the sense of having an autocorrelation function that decays, but which has long memory, in the sense of a rate of decay that is too slow to produce a convergent sum.

The aim of this paper is to re-examine the measurement of persistence in univariate time series models. Of particular focus is the measurement of persistence in nonstationary models that do not fall within the I(1) class. The extension of the model framework uses the decomposition of Cramer (1961), which is a generalization of the decomposition of Wold (1938) for stationary time series. In addition to including the standard I(0) and I(1) models, it enables a variety of alternative nonstationary processes to be accommodated. In deriving the persistence measures for certain of these nonstationary models and demonstrating their behaviour, we are thus using a fairly natural setting in which to highlight issues associated with persistence measurement. That is, the conclusions we draw do not reflect the properties of perverse models but, rather, are indicative of the behaviour of persistence measures in parametric settings not far removed from those conventionally used in univariate time series analysis. We also make note of the memory properties of the alternative models, as based on the summability of the autocorrelation coefficients.

Our discussion is framed entirely in the time domain, but mention should be made of related work in the frequency domain that seeks to model nonstationary and nonlinear processes using spectral methods. The main contributions use the notion of (time dependent) evolutionary spectra; see, for example, Priestley (1981, 11.2). This approach does not seek specifically to examine persistence and relies on the notion that certain processes can be regarded as 'locally stationary' over some neighbourhood around any data point.

The outline of the paper is as follows. In Section 2, the Cramer representation is used to nest a series of models, including models that produce a wider range of nonstationary behaviour than an I(1) process. Section 3 then outlines the persistence measures that are to be considered, namely the long-run impulse response, variance ratio and autocorrelation functions. The nature of the Cramer representation helps clarify exactly what shock the impulse response function refers to in the case of models where more than one random disturbance impinges on the time series. Section 4 explores persistence measurement in each model. For ease of exposition, the results for all models are tabulated measure by measure, with details of the derivations presented in the Appendices. We reveal conflict that arises between different measures, applied to the same model, as well as that arising from the use of a given measure in different models. Precisely which persistence measures are, and are not, time dependent is highlighted for the models considered. Section 5 concludes.

2 A General Representation for Time Series

A convenient representation for any times series, either stationary or nonstationary, is given by Cramer (1961). It is a generalization of the well-known Wold (1938) decomposition, expressing an arbitrary (nondeterministic) series in terms of a linear process, where the coefficients may depend on time and the disturbances, whilst uncorrelated with zero mean, may be heteroskedastic. We use a slight modification of the Cramer representation of a q-dimensional vector time series $\mathbf{y}_t$,

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \mathbf{A}_{t,t-j} \boldsymbol{\varepsilon}_{t-j}, \tag{1}$$

where $\{\boldsymbol{\varepsilon}_t\}$ is a $(q \times 1)$ vector white noise (WN) process with zero mean and covariance matrix, I, and the nonstochastic $(q \times q)$ matrix coefficients $\mathbf{A}_{t,t-j}$ are constructed so as to absorb any heteroskedastic components in the disturbances. The coefficients $\mathbf{A}_{t,t-j}$ are assumed to satisfy

$$\sum_{j=0}^{\infty} \|\mathbf{A}_{t,t-j}\|^2 < \infty \text{ for all } t, \tag{2}$$

where $\|\mathbf{B}\| = \sqrt{tr(\mathbf{B}'\mathbf{B})}$ for any matrix $\mathbf{B}$. The representation of $\mathbf{y}_t$ in (1) subsumes the Wold representation for stationary $\mathbf{y}_t$, namely $\mathbf{A}_{t,t-j} = \mathbf{A}_{t-(t-j)} = \mathbf{A}_j$ for all t.

The representation (1) of a purely non-deterministic process has been used previously in the frequency domain by Tjøstheim (1976) to define a time-varying spectrum. Priestley (1988, 6.4) points out that Tjøstheim's definition is a special case of that of the evolutionary spectrum mentioned above. He also draws attention to the related work of Melard and his co-workers that makes use of Cramer's device in characterizing the stochastic properties of processes; see, for example, Melard (1985) for an accessible reference. This is similar in spirit to the use of (1) in the present paper.

Within the context of the general framework in (1), we investigate a range of models, with particular emphasis given to nonstationary models that extend the I(0)/I(1) paradigm. However, in order to provide a benchmark for the persistence results that relate to these more general nonstationary models, we also consider, as special cases of (1), the I(0) and I(1) specifications. In what follows, we investigate five models in the time domain:

1. the covariance stationary linear process (LP) constructed as the output of a time invariant linear filter of a WN process;

2. the random walk (RW) model constructed as an undiscounted sum of a WN process;

3. the nonstationary fractionally integrated noise (NSFN) model with fractional difference parameter, d, in the range $0.5 \leq d < 1$;

4. the stochastically integrated (SI) model (Harris, McCabe and Leybourne, 2002), which is a multiplicative combination of a RW and a WN process, together with an additive error; and

5. the stochastic unit root (SUR) model (McCabe and Tremayne, 1995, Leybourne, McCabe, and Tremayne, 1996, and Granger and Swanson, 1997) which is a recursive combination of a RW and a WN process.

Model 1 encompasses both I(0) models and stationary I(d) models, with $d < 0.5$, whilst Model 2 is a special case of an I(1) model. Model 3 is a nonstationary I(d) model, whereas Models 4 and 5 are nonstationary but do not fit into either the I(1) or I(d) classes. As shown below, Model 4 is a special case of a nonlinear state space model with a nonstationary state process. For models with nonstationary components we assume that the process begins at some time in the finite past. We now demonstrate the form assumed by both the $\mathbf{A}_{t,t-j}$ and the $\boldsymbol{\varepsilon}_{t-j}$ for each of the different models.

2.1 LP Model

Letting $\mathbf{A}_{t,t-j} = a_j$, a scalar, in (1), with $\sum_{j=0}^{\infty} a_j^2 < \infty$, then

$$y_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \tag{3}$$

provides the usual LP representation for a stationary univariate series, y_t, where $\{\varepsilon_t\}$ is a scalar WN process with constant variance 1. This representation includes the stationary fractional noise model, in which case

$$a_j \propto \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}, \tag{4}$$

with $d < 0.5$.

2.2 RW model

Let $\{\eta_t\}$ be $WN(0, \sigma_\eta^2)$ and set $\{\varepsilon_t\} = \left\{\sigma_\eta^{-1}\eta_t\right\}$. Let $\mathbf{A}_{t,t-j} = a_{t,t-j} + b_j$, where $a_{t,t-j} = \sigma_\eta$ when $t - j > 0$, $a_{t,t-j} = 0$ when $t - j \leq 0$, and b_j is a scalar such that $\sum_{j=0}^{\infty} b_j^2 < \infty$. In this case, (1) gives the representation

$$\begin{aligned} y_t &= a_{t,t}\varepsilon_t + a_{t,t-1}\varepsilon_{t-1} + \cdots + a_{t,1}\varepsilon_1 + \sum_{j=0}^{\infty} b_j \varepsilon_{t-j} \\ &= \sum_{j=0}^{t-1} \sigma_\eta \varepsilon_{t-j} + \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}. \end{aligned}$$

Thus y_t is a RW plus a covariance stationary LP, as in the Beveridge and Nelson (1981) decomposition of an I(1) series. For simplicity we set $b_j = 0$, for all j, so that

$$y_t = \sum_{j=0}^{t-1} \sigma_\eta \varepsilon_{t-j} \tag{5}$$

is a RW with WN disturbances. Alternatively, we can express y_t as

$$\begin{aligned} y_t &= w_t \\ w_t &= w_{t-1} + \eta_t, \end{aligned}$$

where $w_0 = 0$.

2.3 NSFN Model

Letting $\mathbf{A}_{t,t-j} = a_j$, in (1), when $t - j > 0$ and $\mathbf{A}_{t,t-j} = 0$ when $t - j \leq 0$, the NSFN model is represented as

$$y_t = \sum_{j=0}^{t-1} a_j \varepsilon_{t-j}, \tag{6}$$

with a_j as defined in (4), for $0.5 \leq d < 1$ and $\{\varepsilon_t\}$ as defined previously.

2.4 SI model

Let $\{e_t\}$, $\{\eta_t\}$ and $\{v_t\}$ be WN sequences, all three mutually independent of one another, with respective variances σ_e^2, σ_η^2 and σ_v^2. Define, the sequence of random variables

$$\begin{pmatrix} \eta_{t-j} \\ v_{t-j} \sum_{i=1}^{t-j} \eta_i \\ e_{t-j} \end{pmatrix}.$$

It is easy to verify that the above sequence is uncorrelated with covariance matrix

$$\mathbf{V}_{t-j} = \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & (t-j)\sigma_v^2\sigma_\eta^2 & 0 \\ 0 & 0 & \sigma_e^2 \end{bmatrix}$$

at time $t-j$. Set

$$\boldsymbol{\varepsilon}_{t-j} = \begin{pmatrix} \varepsilon_{1,t-j} \\ \varepsilon_{2,t-j} \\ \varepsilon_{3,t-j} \end{pmatrix} = \mathbf{V}_{t-j}^{-1/2} \begin{pmatrix} \eta_{t-j} \\ v_{t-j}\sum_{i=1}^{t-j}\eta_i \\ e_{t-j} \end{pmatrix}, \quad (7)$$

such that $\{\boldsymbol{\varepsilon}_{t-j}\}$ is a WN sequence. Define $\mathbf{A}_{t,t-j} = \mathbf{A}_{t,t-j}^* + \mathbf{B}_{t,t-j}^*$. Here

$$\mathbf{B}_{t,t-j}^* = \begin{bmatrix} 0 & 0 & b_j \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}_{t-j}^{1/2},$$

where b_j is as defined earlier, and

$$\mathbf{A}_{t,t}^* = \begin{bmatrix} \pi & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}_t^{1/2}, \ \mathbf{A}_{t,t-j}^* = \begin{bmatrix} \pi & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}_{t-j}^{1/2}; \ t-j > 0, \ j > 0,$$

with $\mathbf{A}_{t,t-j}^* = \mathbf{0}$ for $t-j \leq 0$, and π an arbitrary constant. Thus we obtain a representation for a 3-dimensional $\mathbf{y}_t$ in (1) of the following form,

$$\mathbf{y}_t = \mathbf{A}_{t,t}^*\boldsymbol{\varepsilon}_t + \mathbf{A}_{t,t-1}^*\boldsymbol{\varepsilon}_{t-1} + \cdots + \mathbf{A}_{t,1}^*\boldsymbol{\varepsilon}_1 + \sum_{j=0}^{\infty}\mathbf{B}_{t,t-j}^*\boldsymbol{\varepsilon}_{t-j}. \quad (8)$$

Focussing on the first component of $\mathbf{y}_t$, we obtain

$$y_{1,t} = y_t = \pi\sum_{i=1}^{t}\eta_i + v_t\sum_{i=1}^{t}\eta_i + \sum_{j=0}^{\infty} b_j e_{t-j}. \quad (9)$$

More compactly, with $u_t = \sum_{j=0}^{\infty} b_j e_{t-j}$, we have

$$\begin{aligned} y_t &= \pi_t w_t + u_t \\ w_t &= w_{t-1} + \eta_t \\ \pi_t &= \pi + v_t, \end{aligned} \quad (10)$$

which is the nonlinear state space model analyzed in Harris, McCabe and Leybourne (2002). For simplicity here and in what follows, we assume that $b_0 = 1$ and $b_j = 0$ for $j > 0$, in which case $u_t = e_t$. In this case, if $\pi = 0$, y_t is an uncorrelated heteroskedastic sequence, as follows from (9). If $\pi \neq 0$, y_t can be viewed as an RW model with a heteroskedastic term. If $\pi = 1$ and $v_t = 0$ for all t, (10) collapses to the Local Level (LL) model of Harvey (1991), which can be viewed as a linear state space model.

2.5 SUR model

Define $\{\eta_t\}$ and $\{v_t\}$ to be WN sequences, independent of one another, with respective variances σ_η^2 and σ_v^2. Set $\alpha_t = 1 + v_t$ and $\xi_t = v_t \sum_{i=1}^{t-2} \left(\prod_{l=i+1}^{t-1} \alpha_l\right) \eta_i$. Note that the random sequence

$$\begin{pmatrix} \eta_{t-j} \\ \xi_{t-j} \end{pmatrix},$$

defined for $t - j > 0$, is uncorrelated with covariance matrix

$$\mathbf{V}_{t-j} = \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\eta^2 \sum_{i=1}^{t-j-2} (\sigma_v^2)^{(t-j-i)} \end{bmatrix}.$$

Let

$$\begin{aligned} \mathbf{A}_{t,t-j} &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{V}_{t-j}^{1/2} \quad \text{for } t - j > 0, \\ &= \mathbf{0} \quad \text{for } t - j \leq 0 \end{aligned} \tag{11}$$

and define

$$\boldsymbol{\varepsilon}_{t-j} = \begin{pmatrix} \varepsilon_{1,t-j} \\ \varepsilon_{2,t-j} \end{pmatrix} = \mathbf{V}_{t-j}^{-1/2} \begin{pmatrix} \eta_{t-j} \\ \xi_{t-j} \end{pmatrix}, \tag{12}$$

such that $\{\boldsymbol{\varepsilon}_{t-j}\}$ is a WN sequence. A representation of a 2-dimensional $\mathbf{y}_t$ in (1) of the following form,

$$\mathbf{y}_t = \mathbf{A}_{t,t}\boldsymbol{\varepsilon}_t + \mathbf{A}_{t,t-1}\boldsymbol{\varepsilon}_{t-1} + \cdots + \mathbf{A}_{t,1}\boldsymbol{\varepsilon}_1,$$

produces a univariate representation for the first component, $y_{1,t} = y_t$, of the form

$$\begin{aligned} y_t &= \alpha_t y_{t-1} + \eta_t \\ \alpha_t &= 1 + v_t, \end{aligned} \tag{13}$$

which is the more familiar form of the SUR model.

3 Persistence Measures for the General Cramer Representation

The focus of the paper is on illustrating the behaviour of standard persistence measures in the range of models delineated above. It will be seen that the outcomes regarding persistence can be both counter-intuitive and indeed conflicting across measures. The measures considered are: the long-run impulse response; the long-run value of the variance ratio; and the long-run autocorrelation. We begin by expressing each measure using the general, multivariate representation in (1) above. In so doing, we make clear which white noise innovations are driving each measure. This is particularly important when it comes to the interpretation of the impulse response function in those models where more than one random innovation is present. As is shown below, it follows from (1) and (2) that *any* model nested therein, whether stationary or nonstationary, has a long-run autocorrelation function that decays to zero, thus indicating a lack of persistence. In the following section we look at the form of each measure for the five models of Section 2.

When dealing with nonstationary models, it becomes necessary to be precise about the way in which persistence measures are defined and interpreted. This is because of: 1) the lack of invariance of persistence measures to the lead/lag distinction; and 2) the dependence of such measures on t. With regard to the first point, we define all measures in lead form so that persistence is assessed by the extent to which shocks at time t impact on the value of $\mathbf{y}_{t+k}$ as $k \rightarrow \infty$. Given that all nonstationary processes are assumed to begin at some point in the finite past, using leads rather than lags is the only way limiting (persistence) behaviour can be gauged. With regard to the second point, we consider the limiting behaviour of all measures over k for a fixed value of t. In this way, the interpretation of persistence measures is clear: given a shock at time period t, the measures gauge the extent of the effect on the value of $\mathbf{y}_{t+k}$ k periods into the future. For any of the three measures, a zero value in the long-run implies a lack of persistence, whilst a non-zero value indicates the converse. The magnitude of the long-run value is indicative of the extent of the persistence in the series, although the precise interpretation of that value varies with the measure.

3.1 Long-run Impulse Response

An impulse response attempts to measure the impact of a unit shock to the innovation vector at time t on $\mathbf{y}_{t+k}$ for some $k > 0$. From (1)

$$\mathbf{y}_{t+k} = \sum_{j=0}^{\infty} \mathbf{A}_{t+k,t+k-j} \boldsymbol{\varepsilon}_{t+k-j} \tag{14}$$

and we define the k-th matrix of impulse responses as

$$\begin{aligned} \mathbf{IR}_{k,t} &= \frac{\partial \mathbf{y}_{t+k}}{\partial \boldsymbol{\varepsilon}_t^{\prime}} \\ &= \mathbf{A}_{t+k,t}. \end{aligned}$$

In general, this quantity may be a function of both t and k. The long-run matrix of impulse responses is defined as

$$\begin{aligned}\mathbf{IR}_t &= \lim_{k\to\infty}\frac{\partial \mathbf{y}_{t+k}}{\partial \boldsymbol{\varepsilon}_t^{'}} \\ &= \lim_{k\to\infty}\mathbf{A}_{t+k,t},\end{aligned}$$

for a fixed value of t.

When the object of interest is the first component of $\mathbf{y}_{t+k}$, denoted by $y_{1,t+k} = y_{t+k}$, $\mathbf{IR}_t$ is a $(q \times 1)$ *vector* of impulse responses

$$\begin{aligned}\mathbf{IR}_t &= \lim_{k\to\infty}\frac{\partial y_{t+k}}{\partial \boldsymbol{\varepsilon}_t} \\ &= \lim_{k\to\infty}\mathbf{a}_{t+k,t}^{'},\end{aligned} \tag{15}$$

where $\mathbf{a}_{t+k,t}$ denotes the first row of $\mathbf{A}_{t+k,t}$.

3.2 Long-run Variance Ratio

By analogy with the scalar version as used in Cochrane (1988) and Campbell and Mankiw (1989), amongst others, the variance ratio is defined as

$$\begin{aligned}V_{k,t} &= \frac{1}{k}\times\frac{tr\{var(\mathbf{y}_{t+k}-\mathbf{y}_t)\}}{tr\{var(\mathbf{y}_{t+1}-\mathbf{y}_t)\}} \\ &= \frac{1}{k}\times\frac{tr\{var(\mathbf{y}_{t+k})+var(\mathbf{y}_t)-2cov(\mathbf{y}_{t+k},\mathbf{y}_t)\}}{tr\{var(\mathbf{y}_{t+1})+var(\mathbf{y}_t)-2cov(\mathbf{y}_{t+1},\mathbf{y}_t)\}}.\end{aligned} \tag{16}$$

Using the respective Cramer representations of $\mathbf{y}_{t+k}$ and $\mathbf{y}_t$ in (14) and (1), $V_{k,t}$ can be expressed as

$$\begin{aligned}V_{k,t} = \frac{1}{k}\times tr&\left[\left(\sum_{j=0}^{\infty}\mathbf{A}_{t+k,t+k-j}\mathbf{A}_{t+k,t+k-j}^{'}\right)+\left(\sum_{j=0}^{\infty}\mathbf{A}_{t,t-j}\mathbf{A}_{t,t-j}^{'}\right)-2\sum_{j=0}^{\infty}\mathbf{A}_{t+k,t-j}\mathbf{A}_{t,t-j}^{'}\right] \\ \times tr&\left[\left(\sum_{j=0}^{\infty}\mathbf{A}_{t+1,t+1-j}\mathbf{A}_{t+1,t+1-j}^{'}\right)+\left(\sum_{j=0}^{\infty}\mathbf{A}_{t,t-j}\mathbf{A}_{t,t-j}^{'}\right)-2\sum_{j=0}^{\infty}\mathbf{A}_{t+1,t-j}\mathbf{A}_{t,t-j}^{'}\right]^{-1}.\end{aligned}$$

In general, $V_{k,t}$ is a function of both t and k. For fixed t the long-run variance ratio is defined as

$$V_t = \lim_{k\to\infty}V_{k,t}. \tag{17}$$

When the objects of interest are the first elements of $\mathbf{y}_{t+k}$ and $\mathbf{y}_t$, y_{t+k} and y_t respectively, $V_{k,t}$ becomes

$$V_{k,t} = \frac{1}{k}\times\frac{var(y_{t+k}-y_t)}{var(y_{t+1}-y_t)}. \tag{18}$$

With reference to the version of $V_{k,t}$ in (18), Cochrane (1988) uses the method of Beveridge and Nelson (1981) to decompose y_t into permanent and temporary components and to show that, in the context of I(0)/I(1) models, V_t measures the proportion of variation in Δy_t attributable to variation in innovations of the permanent component. For an I(0) process, this proportion is equal to zero, whilst for a random walk, it is equal to one. For a general I(1) process, the proportion may be less than or greater than unity, depending on the nature of the dynamics in the temporary component, but is always greater than zero. In the more general context of (1) these interpretations of V_t no longer necessarily hold.

3.3 Long-run Autocorrelation

Let

$$\mathbf{D}_t = [diag\{var(\mathbf{y}_t)\}]^{1/2} = \left[diag\{\sum_{j=0}^{\infty} \mathbf{A}_{t,t-j}\mathbf{A}'_{t,t-j}\}\right]^{1/2}.$$

That is, $\mathbf{D}_t$ is a diagonal matrix with the standard deviations of the elements of $\mathbf{y}_t$ along the main diagonal. The k-th autocorrelation matrix between $\mathbf{y}_{t+k}$ and $\mathbf{y}_t$ is defined as

$$\mathbf{\Omega}_{k,t} = \mathbf{D}_{t+k}^{-1} cov(\mathbf{y}_{t+k}, \mathbf{y}_t)\mathbf{D}_t^{-1}.$$

Using (1), $\mathbf{\Omega}_{k,t}$ can be expressed as

$$\begin{aligned}\mathbf{\Omega}_{k,t} &= \mathbf{D}_{t+k}^{-1} E\left[\left(\sum_{j=0}^{\infty} \mathbf{A}_{t+k,t+k-j}\boldsymbol{\varepsilon}_{t+k-j}\right)\left(\sum_{j=0}^{\infty} \mathbf{A}_{t,t-j}\boldsymbol{\varepsilon}_{t-j}\right)'\right]\mathbf{D}_t^{-1} \\ &= \mathbf{D}_{t+k}^{-1}\left[\sum_{j=0}^{\infty} \mathbf{A}_{t+k,t-j}\mathbf{A}'_{t,t-j}\right]\mathbf{D}_t^{-1}.\end{aligned}$$

In general, $\mathbf{\Omega}_{k,t}$ is a function of both t and k. The long-run autocorrelation matrix is defined as

$$\mathbf{\Omega}_t = \lim_{k\to\infty} \mathbf{\Omega}_{k,t}, \tag{20}$$

for a fixed value of t.

For *all* processes that satisfy the representation in (1) the limit in (20) is zero. To see this write

$$Cov(\mathbf{y}_{t+k}, \mathbf{y}_t) = \sum_{j=0}^{\infty} \mathbf{A}_{t+k,t-j}\mathbf{A}'_{t,t-j},$$

such that

$$
\begin{aligned}
\|Cov(\mathbf{y}_{t+k}, \mathbf{y}_t)\| &= \left\| \sum_{j=0}^{\infty} \mathbf{A}_{t+k,t-j} \mathbf{A}'_{t,t-j} \right\| \\
&\leq \sum_{j=0}^{\infty} \left\| \mathbf{A}_{t+k,t-j} \mathbf{A}'_{t,t-j} \right\| \\
&= \sum_{j=0}^{\infty} \left\| \mathbf{A}_{t+k,t-j} \right\| \left\| \mathbf{A}_{t,t-j} \right\|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sum_{j=0}^{\infty} \left\| \mathbf{A}_{t+k,t-j} \right\|^2 \sum_{j=0}^{\infty} \left\| \mathbf{A}_{t,t-j} \right\|^2 \\
&= \sum_{j=0}^{\infty} \left\| \mathbf{A}_{t,t-k-j} \right\|^2 \times const \\
&= \sum_{j=k}^{\infty} \left\| \mathbf{A}_{t,t-j} \right\|^2 \times const \\
&\rightarrow 0 \text{ for all t as } k \rightarrow \infty.
\end{aligned}
$$

Since the diagonal elements of $\mathbf{D}_t$ are bounded away from zero, it follows that

$$\mathbf{\Omega}_t = \lim_{k \to \infty} \mathbf{\Omega}_{k,t} = \mathbf{0} \tag{21}$$

for all t. Thus, for any model in the Cramer class, the correlation between any corresponding elements of $\mathbf{y}_{t+k}$ and $\mathbf{y}_t$ approaches zero with k, for any value of t, indicating that the series is not persistent. When the objects of interest are the first elements of $\mathbf{y}_{t+k}$ and $\mathbf{y}_t$, y_{t+k} and y_t respectively, the kth order autocorrelation coefficient is denoted by $\rho_{k,t}$ and is given by

$$\rho_{k,t} = \frac{cov(y_{t+k}, y_t)}{\sqrt{var(y_{t+k})var(y_t)}} \tag{22}$$

and the result in (21) can be simply stated as

$$\rho_t = \lim_{k \to \infty} \rho_{k,t} = 0. \tag{23}$$

Note that this definition of the limiting behaviour of the autocorrelation coefficient differs from the alternative that is sometimes invoked with reference to the $I(1)$ model, whereby for fixed k, $\rho_{k,t} \rightarrow 1$ as $t \rightarrow \infty$.

Table 1: Impulse Response Functions

Model	**Eqn.**	$\mathbf{IR}_{k,t}$	$\mathbf{IR}_t = \lim_{k\to\infty} \mathbf{IR}_{k,t}$
LP	(3)	a_k	0
RW	(5)	σ_η	σ_η
NSFN	(6)	a_k	0
SI	(10)	$\frac{\partial y_{t+k}}{\partial \varepsilon_{1,t}} = \sigma_\eta \pi$	$\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_{1,t}} = \sigma_\eta \pi$
		$\frac{\partial y_{t+k}}{\partial \varepsilon_{2,t}} = 0$	$\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_{2,t}} = 0$
		$\frac{\partial y_{t+k}}{\partial \varepsilon_{3,t}} = \sigma_e b_k$	$\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_{3,t}} = 0$
SUR	(13)	$\frac{\partial y_{t+k}}{\partial \varepsilon_{1,t}} = \sigma_\eta$	$\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_{1,t}} = \sigma_\eta$
		$\frac{\partial y_{t+k}}{\partial \varepsilon_{2,t}} = \sqrt{\sigma_\eta^2 \sum_{i=1}^{t-2} \left(\sigma_v^2\right)^{(t-i)}}$	$\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_{2,t}} = \sqrt{\sigma_\eta^2 \sum_{i=1}^{t-2} \left(\sigma_v^2\right)^{(t-i)}}$

4 Persistence Measures for the Five Models

In this section we explore the behaviour of the various measures introduced above for Models 1 to 5 described in Section 2. The emphasis is on the behaviour of the population quantities and their interpretation. For ease of exposition we tabulate the results for all five models, measure by measure. The detailed derivations of the results are provided in the Appendices. We include in the tables the equation number associated with the most convenient representation of each model.

We have already demonstrated, in Section 3.3, that the limiting value of the autocorrelation function is equal to zero for *all* models subsumed in (1). Hence, the only issue left to address with respect to that particular measure of persistence is the *rate* of convergence to zero and the consequent conclusions regarding memory. For the two other persistence measures, there is no general result for models represented by (1); the specific models under consideration have both differing limiting values for the relevant persistence measures and differing rates of convergence to those values. In several cases, the long-run impulse response and long-run variance ratio functions provide a measure of persistence which conflicts with that provided by the long-run autocorrelation function.

4.1 Long-run Impulse Response

For the LP, RW and NSFN models the single source of error means that the impulse response function, $\mathbf{IR}_t$ in (15), reduces to a scalar quantity measuring the response of y_{t+k} to a unit shock in the error term ε_t; that is,

$$IR_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_t}.$$

On the other hand, for the SUR model, $\mathbf{IR}_t$ is a two-dimensional vector measuring the response of y_{t+k} to unit shocks in $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, as defined in (12). For the SI model $\mathbf{IR}_t$ is a three-dimensional vector measuring the response of y_{t+k} to unit shocks in $\varepsilon_{1,t}$, $\varepsilon_{2,t}$ and $\varepsilon_{3,t}$ defined in (7). In Table 1, we present the results for all models, with the appropriate number of impulse responses reported. We give the k-dependent form of each measure, as well as the limit of each measure as $k \to \infty$.

As indicated by the values recorded in Table 1, both the LP and NSFN models are nonpersistent according to this measure. In contrast, the RW and SUR models are clearly persistent, with unit shocks to the relevant innovations in the current period continuing to affect the level of y_{t+k} in the infinite future, that is, as $k \to \infty$. In particular, despite the additional noise incorporated in the random coefficient of the SUR model, the long-run response to a shock in the standardized random walk innovation, $\varepsilon_{1,t} = \sigma_\eta^{-1}\eta_t$, is equivalent to the response induced by the parallel shock in the RW model. A unit shock to the second random component of the model $\varepsilon_{2,t} = \left(\sigma_\eta^2 \sum_{i=1}^{t-2} \left(\sigma_v^2\right)^{(t-i)}\right)^{-1/2} \left(v_t \sum_{i=1}^{t-2} \left(\prod_{l=i+1}^{t-1} \alpha_l\right) \eta_i\right)$ produces a change in y_{t+k} equal to $\left(\sigma_\eta^2 \sum_{i=1}^{t-2} \left(\sigma_v^2\right)^{(t-i)}\right)^{1/2}$, which is a function of t, but not k.

The results for the SI model also indicate a non-zero long-run response to a shock in the standardized random walk innovation, $\varepsilon_{1,t}$, as long as $\pi \neq 0$. To this extent, the persistence properties of the SI and SUR models mimic those of the RW model. However, the long-run impact of a shock to the second and third random components in the SI model, namely $\varepsilon_{2,t} = \left(t\sigma_v^2\sigma_\eta^2 + \sigma_e^2\right)^{-1/2} \left(v_t \sum_{i=1}^{t} \eta_i + e_t\right)$ and $\varepsilon_{3,t} = \sigma_e^{-1} e_t$ respectively, is zero. Also, if $\pi = 0$, the SI model exhibits no long-run response to a shock in any of the random innovations in the model, despite the presence of the RW component.

Note that only in the case of the SUR model does an impulse response coefficient depend on t. That is, despite being nonstationary, the RW, NSFN and SI models have impulse response coefficients that do not depend on the period of time at which the coefficient is evaluated. As is clear from the expression reported in the final column in Table 1, the long-run impulse response function with respect to $\varepsilon_{2,t}$ for the SUR model is not only a function of t but also explodes as $t \to \infty$.

4.2 Long-run Variance Ratio

The variance ratio results are reported in Table 2. When reporting the value of $V_{k,t}$ for the LP model, we refer to the variance of y_t as γ_0, and the kth autocovariance as γ_k for $k \geq 1$. As is evident, the long-run variance ratio results are consistent with the long-run impulse response results for the LP, RW and NSFN models. That is, according to both measures, neither the LP nor NSFN model is persistent, whilst the RW is persistent. For the SI model the additional noise, relative to the RW model, reduces the measure of long-run persistence to something less than unity. However, the variance ratio is still non-zero for this model, as well as being dependent on t, for any value of π. Note that when $\pi = 0$, $V_t \neq 0$, thus indicating persistence, although it will be recalled that the SI process is actually uncorrelated in this case. Whilst the impulse response function for this model indicates persistence to only one shock, $\varepsilon_{1,t}$, and then only for $\pi \neq 0$, the long-run variance ratio is unambiguous in its finding of persistence for any value of π.

In the case of the SUR model, the addition of noise to the RW component serves to *increase* rather than decrease the degree of persistence as indicated by this particular measure. Indeed, the long-run variance ratio is infinite for this model, with shocks accumulating at an increasing rate with k.

4.3 Long-run Autocorrelation

As demonstrated in Section 3.3, all models that fall within the Cramer class have long-run autocorrelation coefficients of zero. For the five specific models under consideration, the precise form of the autocorrelation function, for finite k, is recorded in Table 3, along with the zero long-run value. The final column in the table reports the summability of the autocorrelation coefficients as a measure of the rate of decline towards zero of each function. In order to cater for the nonstationary models which, by assumption, begin at a point in the finite past, the summation extends only into the infinite future. As noted in the Introduction, the typical definition of a long memory process is one for which the autocorrelations are not absolutely summable. Summability implies that the process has short, or perhaps, intermediate memory (see Brockwell and Davis, 1991). If persistence is

Table 2: Variance Ratio Functions

Model	**Eqn.**	$V_{k,t}$	$V_t = \lim_{k\to\infty} V_{k,t}$
LP	(3)	$\frac{1}{k}\frac{2\gamma_0-2\gamma_k}{2\gamma_0-2\gamma_1}$	0
RW	(5)	$\frac{1}{k}\frac{k\sigma_\varepsilon^2}{\sigma_\varepsilon^2}=1$	1
NSFN	(6)	$\frac{1}{k}\frac{\sum_{j=0}^{t+k-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}+\sum_{j=0}^{t-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}-2\sum_{j=0}^{t-1}\frac{\Gamma(k+j+d)}{\Gamma(k+j+1)}\frac{\Gamma(j+d)}{\Gamma(j+1)}}{\sum_{j=0}^{t}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}+\sum_{j=0}^{t-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}-2\sum_{j=0}^{t-1}\frac{\Gamma(1+j+d)}{\Gamma(2+j)}\frac{\Gamma(j+d)}{\Gamma(j+1)}}$	0
SI	(10)	$\frac{1}{k}\left(\frac{c+k\sigma_\eta^2[\pi^2+\sigma_v^2]}{c+\sigma_\eta^2[\pi^2+\sigma_v^2]}\right)$; $c=2\sigma_e^2+2t\sigma_\eta^2\sigma_v^2$	$0<\left(\frac{\sigma_\eta^2[\pi^2+\sigma_v^2]}{c+\sigma_\eta^2[\pi^2+\sigma_v^2]}\right)<1$
SUR	(13)	$\frac{1}{k}\frac{(1+\sigma_v^2)^k-1}{\sigma_v^2}$	∞

defined in terms of the value of the long-run autocorrelation coefficient alone, then it is quite feasible that a series may be nonpersistent yet at the same time exhibit long memory. As is clear from the results reported in Table 3, this is the case for the RW, NSFN and SI models. In the case of the SUR model, the fast rate of decline in $\rho_{k,t}$, for fixed t, induces summability, indicating that the SUR model is a short-memory process. The first line of Table 3 shows that the summability of the $\rho_{k,t}$ in the LP model depends on the particular structure of γ_k. For example, the LP model covers the stationary fractionally integrated model ($0 < d < 0.5$), for which the rate of decline in γ_k is not rapid enough to ensure summability. On the other hand, for $d \leq 0$, the autocorrelations are summable.

4.4 Discussion

Table 4 summarizes the persistence results for the three measures used; P indicates persistence and NP lack thereof. In the long-run impulse response column, the nth entry in any row corresponds to the nth element of the $(q \times 1)$ innovation vector $\boldsymbol{\varepsilon}_{t-j}$ in the Cramer decomposition (1). The fact that the autocorrelation function indicates a lack of persistence for all models under consideration obviously conflicts with the other persistence measures for the RW, SI and SUR models. Only for the LP and NSFN models are the results associated with all three measures consistent in indicating that the models are not persistent. For the RW, SI and SUR models, the long-run variance ratio provides the opposite qualitative result

Table 3: Autocorrelation Functions

Model	**Eqn.**	$\rho_{k,t}$	$\rho_t = \lim_{k\to\infty} \rho_{k,t}$	$\sum_{k=0}^{\infty} \left\|\rho_{k,t}\right\|$
LP	(3)	$\frac{\gamma_k}{\gamma_0}$	0	$\left.\begin{matrix}\to \\ <\end{matrix}\right\}\infty$
RW	(5)	$\sqrt{\frac{t}{t+k}}$	0	$\to\infty$
NSFN	(6)	$\frac{\sum_{j=0}^{t-1} \frac{\Gamma(k+j+d)}{\Gamma(k+j+1)} \frac{\Gamma(j+d)}{\Gamma(j+1)}}{\sqrt{\left(\sum_{j=0}^{t+k-1} \frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}\right)\left(\sum_{j=0}^{t-1} \frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}\right)}}$	0	$\to\infty$
SI	(10)	$\frac{t\sigma_\eta^2\pi^2}{\sqrt{((t+k)\sigma_\eta^2[\pi^2+\sigma_v^2]+\sigma_e^2)(t\sigma_\eta^2[\pi^2+\sigma_v^2]+\sigma_e^2)}}$	0	$\to\infty$
SUR	(13)	$\sqrt{\frac{(1+\sigma_v^2)^t-1}{(1+\sigma_v^2)^{t+k}-1}}$	0	$<\infty$

Table 4: Summary of Persistence Measures

P = Persistent; NP = Nonpersistent.

Model	**Eqn.**	$\mathbf{IR}_t$			V_t	ρ_t
LP	(3)	NP			NP	NP
RW	(5)	P			P	NP
NSFN	(6)	NP			NP	NP
			$\pi = 0$	$\pi \neq 0$		
SI	(10)	$\mathbf{IR}_t[1]$	NP	P	P	NP
		$\mathbf{IR}_t[2]$	NP	NP		
		$\mathbf{IR}_t[3]$	NP	NP		
SUR	(13)	$\mathbf{IR}_t[1]$	P		P	NP
		$\mathbf{IR}_t[2]$	P			

to the autocorrelation function in all three cases, whilst the long-run impulse response function conflicts with the autocorrelation function for the RW and SUR models. The impulse response function for the SI model is dependent upon both the shock being considered and the value of the parameter π. Hence the results associated with this measure may, or may not, be in conflict with those arising from the autocorrelation function.

In the SI and SUR models V_t and $\mathbf{IR}_t$ both depend on t and, thus, are not invariant measures of persistence. As such, interpretation of these measures is not clear-cut, with the time invariant underpinning of descriptive statistics no longer applicable, a point further amplified by noting that V_t in the SUR model is infinite in k. Indeed, as pointed out by Phillips (2001), there have until recently been no usable descriptive statistics tools for nonstationary data. The embryonic results on spatial densities and sojourn distributions reported by Phillips (2001) for nonstationary models like a random walk with drift may, in due course, provide a feasible methodology for dealing with time-dependent persistence measure such as those encountered here.

5 Conclusions

The paper has documented the behaviour of persistence measures in a series of models nested in the general representation of Cramer (1961). The key message is that all of the traditional measures of persistence can lead to ambiguous results, even in the standard RW model, with there being particular problems associated with the use of the measures in more general nonstationary settings. Specifically, there is conflict between the three measures considered for all models other than the LP and NSFN specifications, where all three measures equal zero, indicating a lack of persistence. For the RW model, the impulse response and variance ratio measures are consistent one with the other, the finite value assumed by both measures being conventionally interpreted as indicating a high degree of persistence for this model. In contrast, the autocorrelation function declines to zero as $k \to \infty$, indicating a lack of persistence.

For the SUR model, the impulse response and variance ratio measures not only conflict with the autocorrelation measure, but also with one another. The impulse response coefficients indicate a non-zero but finite degree of persistence (for fixed t at least) whereas the variance ratio indicates a degree of persistence that is unbounded in k. Further, the evidence in favour of persistence is in stark contrast not only with the zero long-run autocorrelation coefficient, but with the rapid geometric decline in the autocorrelation function.

For the SI model, the impulse response and variance ratio measures are consistent with one another when considering the impulse response coefficient on the RW innovation, at least for $\pi \neq 0$. Both, however, conflict with the autocorrelation measure, in indicating persistence. When $\pi = 0$, both the impulse response and autocorrelation measures indicate a lack of persistence. However, according to the variance ratio measure, the model still maintains persistence, since $V_t \neq 0$.

Appendix 1: Impulse Response Functions

The derivations of the results in the column headed $\mathbf{IR}_t$ in Table 1 are provided here, according to the order in which they appear in the table. All results are special cases of the general expression for $\mathbf{IR}_t$ in (15).

LP

For the LP model in (3), with the scalar WN innovation process $\{\varepsilon_t\}$, the kth order coefficient in (14) is a scalar given by $\mathbf{A}_{t+k,t} = a_k$. Given that $\sum_{j=0}^{\infty} a_j^2 < \infty$, it follows that $\mathbf{IR}_t$ is a scalar given by

$$\mathbf{IR}_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_t} = \lim_{k\to\infty} a_k = 0.$$

RW

For the RW model in (5), with the scalar WN innovation process $\{\varepsilon_t\}$, the kth order coefficient in (14) is a scalar given by $\mathbf{A}_{t+k,t} = \sigma_\eta$, from which it follows that $\mathbf{IR}_t$ is a scalar given by

$$\mathbf{IR}_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_t} = \lim_{k\to\infty} \sigma_\eta = \sigma_\eta.$$

NSFN

For the NSFN model in (6), with the scalar WN innovation process $\{\varepsilon_t\}$, the kth order coefficient in (14) is a scalar given by $\mathbf{A}_{t+k,t} = a_k$, with a_k as defined in (4). Since $\lim_{k\to\infty} a_k = 0$ for $d < 1$ (see Baillie, 1996) it follows that $\mathbf{IR}_t$ is a scalar given by

$$\mathbf{IR}_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \varepsilon_t} = \lim_{k\to\infty} a_k = 0.$$

SI

Using the definition of the relevant vectors and matrices in (8), $\mathbf{IR}_t$ is a three-dimensional vector given by

$$\mathbf{IR}_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \boldsymbol{\varepsilon}_t} = \lim_{k\to\infty} \begin{bmatrix} \sigma_\eta \pi \\ 0 \\ \sigma_e b_j \end{bmatrix} = \begin{bmatrix} \sigma_\eta \pi \\ 0 \\ 0 \end{bmatrix},$$

where the last term in the vector follows from the assumption that $\sum_{j=0}^{\infty} b_j^2 < \infty$.

SUR

Using the definition of the 2-dimensional WN sequence $\boldsymbol{\varepsilon}_{t-j}$ in (12), and the definition of $\mathbf{A}_{t,t-j}$ in (11), $\mathbf{IR}_t$ is a two-dimensional vector given by

$$\mathbf{IR}_t = \lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \boldsymbol{\varepsilon}_t} = \lim_{k\to\infty} \begin{bmatrix} \sigma_\eta \\ \sqrt{\sigma_\eta^2 \sum_{i=1}^{t-2} (\sigma_v^2)^{(t-i)}} \end{bmatrix} = \begin{bmatrix} \sigma_\eta \\ \sqrt{\sigma_\eta^2 \sum_{i=1}^{t-2} (\sigma_v^2)^{(t-i)}} \end{bmatrix}.$$

Appendix 2: Variance Ratio Functions

The derivations of the results in the column headed V_t in Table 2 are provided here, according to the order in which they appear in the table. All results are special cases of the general expression for V_t in (17), with $V_{k,t}$ as defined in (18).

LP

For the covariance stationary LP process in (3),

$$var(y_{t+k}) = var(y_{t+1}) = var(y_t) = \gamma_0,$$

$$cov(y_{t+1}, y_t) = \gamma_1$$

and

$$\lim_{k\to\infty} cov(y_{t+k}, y_t) = \lim_{k\to\infty} \gamma_k = 0.$$

Hence,

$$V_t = \lim_{k\to\infty} \left\{ \frac{1}{k} \times \frac{var(y_{t+k} - y_t)}{var(y_{t+1} - y_t)} \right\} = \lim_{k\to\infty} \left\{ \frac{1}{k} \times \frac{2(\gamma_0 - \gamma_1)}{2(\gamma_0 - \gamma_k)} \right\} = 0.$$

RW

For the RW process in (5),

$$var(y_{t+1}-y_t)=var(\sigma_\eta\varepsilon_{t+1})=\sigma_\eta^2$$

and

$$var(y_{t+k}-y_t)=var(\sum_{i=1}^{k}\sigma_\eta\varepsilon_{t+i})=\sigma_\eta^2k.$$

Hence,

$$V_t = \lim_{k\to\infty}\left\{\frac{1}{k}\times\frac{\sigma_\eta^2k}{\sigma_\eta^2}\right\}=\lim_{k\to\infty}\{1\}=1.$$

NSFN

For the NSFN model in (6), with a_j as defined in (4),

$$var(y_{t+k})=var(\sum_{j=0}^{t+k-1}a_j\varepsilon_{t-j})=\sum_{j=0}^{t+k-1}a_j^2, \tag{24}$$

for $k\geq 0$ and

$$cov(y_{t+k},y_t)=\sum_{j=0}^{t-1}a_{j+k}a_j, \tag{25}$$

for $k\geq 1$. Hence,

$$\begin{aligned}V_t &= \lim_{k\to\infty}\left\{\frac{\sum_{j=0}^{t+k-1}a_j^2+\sum_{j=0}^{t-1}a_j^2-2\sum_{j=0}^{t-1}a_{j+k}a_j}{\sum_{j=0}^{t}a_j^2+\sum_{j=0}^{t-1}a_j^2-2\sum_{j=0}^{t-1}a_{j+1}a_j}\right\}\\ &= \lim_{k\to\infty}\left\{\frac{1}{k}\times\frac{\sum_{j=0}^{t+k-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}+\sum_{j=0}^{t-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}-2\sum_{j=0}^{t-1}\frac{\Gamma(k+j+d)}{\Gamma(k+j+1)}\frac{\Gamma(j+d)}{\Gamma(j+1)}}{\sum_{j=0}^{t}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}+\sum_{j=0}^{t-1}\frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}-2\sum_{j=0}^{t-1}\frac{\Gamma(1+j+d)}{\Gamma(2+j)}\frac{\Gamma(j+d)}{\Gamma(j+1)}}\right\}\\ &= 0.\end{aligned}$$

SI

For the SI model in (10),

$$var(y_{t+k})=(t+k)\sigma_\eta^2[\pi^2+\sigma_v^2]+\sigma_e^2, \tag{26}$$

for $k \geq 0$ and

$$cov(y_{t+k}, y_t) = t\sigma_\eta^2 \pi^2, \tag{27}$$

for $k \geq 1$. Using $c = 2\sigma_e^2 + 2t\sigma_\eta^2\sigma_v^2$,

$$V_t = \lim_{k\to\infty} \left\{ \frac{1}{k} \left(\frac{c + k\sigma_\eta^2[\pi^2 + \sigma_v^2]}{c + \sigma_\eta^2[\pi^2 + \sigma_v^2]} \right) \right\} = \frac{\sigma_\eta^2[\pi^2 + \sigma_v^2]}{c + \sigma_\eta^2[\pi^2 + \sigma_v^2]}.$$

Since $c > 0$ it follows that $0 < V_t < 1$.

SUR

For the SUR model in (13),

$$var(y_{t+k}) = \frac{\sigma_\eta^2}{\sigma_v^2} \left[(1 + \sigma_v^2)^{t+k} - 1 \right], \tag{28}$$

for $k \geq 0$ and

$$cov(y_{t+k}, y_t) = \frac{\sigma_\eta^2}{\sigma_v^2} \left[(1 + \sigma_v^2)^{t} - 1 \right], \tag{29}$$

for $k \geq 1$. Hence,

$$V_t = \lim_{k\to\infty} \left\{ \frac{1}{k} \left(\frac{(1 + \sigma_v^2)^k - 1}{\sigma_v^2} \right) \right\} = \infty.$$

Appendix 3: Autocorrelation Functions

The derivations of the results in the column headed ρ_t in Table 3 are provided here, according to the order in which they appear in the table. All results are special cases of the general expression for ρ_t in (23), with $\rho_{k,t}$ as defined in (22).

LP

For the covariance stationary LP process in (3),

$$\rho_t = \lim_{k\to\infty} \left\{ \frac{\gamma_k}{\gamma_0} \right\} = 0.$$

since $\lim_{k\to\infty} \gamma_k = 0$.

RW

For the RW process in (5),

$$var(y_{t+k}) = \sigma_\eta^2(t+k),$$

for $k \geq 0$ and

$$cov(y_{t+k}, y_t) = \sigma_\eta^2 t,$$

for $k \geq 1$. Hence,

$$\rho_t = \lim_{k\to\infty}\left\{\frac{\sigma_\eta^2 t}{\sqrt{\sigma_\eta^2(t+k)\sigma_\eta^2 t}}\right\} = \lim_{k\to\infty}\left\{\sqrt{\frac{t}{(t+k)}}\right\} = 0.$$

<u>NSFN</u>

For the NSFN model in (6), with a_j as defined in (4), $var(y_{t+k})$ and $cov(y_{t+k}, y_t)$ are as defined in (24) and (25) respectively.

$$\begin{aligned}
\rho_t &= \lim_{k\to\infty}\left\{\frac{\sum_{j=0}^{t-1} a_{j+k}a_j}{\sqrt{\left(\sum_{j=0}^{t+k-1} a_j^2\right)\left(\sum_{j=0}^{t-1} a_j^2\right)}}\right\} \\
&= \lim_{k\to\infty}\left\{\frac{\sum_{j=0}^{t-1} \frac{\Gamma(k+j+d)}{\Gamma(k+j+1)}\frac{\Gamma(j+d)}{\Gamma(j+1)}}{\sqrt{\left(\sum_{j=0}^{t+k-1} \frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}\right)\left(\sum_{j=0}^{t-1} \frac{\Gamma^2(j+d)}{\Gamma^2(j+1)}\right)}}\right\} \\
&= 0.
\end{aligned}$$

<u>SI</u>

For the SI model in (10), $var(y_{t+k})$ and $cov(y_{t+k}, y_t)$ are as defined in (26) and (27) respectively. Hence,

$$\rho_t = \lim_{k\to\infty}\left\{\frac{t\sigma_\eta^2\pi^2}{\sqrt{((t+k)\sigma_\eta^2[\pi^2+\sigma_v^2]+\sigma_e^2)(t\sigma_\eta^2[\pi^2+\sigma_v^2]+\sigma_e^2)}}\right\} = 0.$$

<u>SUR</u>

For the SUR model in (13), $var(y_{t+k})$ and $cov(y_{t+k}, y_t)$ are as defined in (28) and (29). Hence,

$$\rho_t = \lim_{k\to\infty}\left\{\sqrt{\frac{(1+\sigma_v^2)^t-1}{(1+\sigma_v^2)^{t+k}-1}}\right\} = 0.$$

References

[1] BAILLIE, R.T. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5-59.

[2] 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, 151-74.

[3] BROCKWELL, P.J. and DAVIS, R.A. (1991) *Time Series: Theory and Methods,* Second Edition, Springer Verlag, New York.

[4] CAMPBELL, J.Y. and MANKIW, N.G. (1987) Are output fluctuations transitory? *Journal of Political Economy* 102, 857-80.

[5] CAMPBELL, J.Y. and MANKIW, N.G. (1989) International evidence on the persistence of economic fluctuations. *Journal of Monetary Economics* 23, 319-33.

[6] COCHRANE (1988) How big is the random walk in GNP? *Journal of Political Economy* 96, 893-920.

[7] CRAMER, H. (1961) On some classes of nonstationary stochastic processes. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability,* Berkeley and Los Angeles, University of Los Angeles Press, 57-77.

[8] GRANGER, C.W.J. and SWANSON, N.R. (1997) An introduction to stochastic unit-root processes. *Journal of Econometrics* 80, 35-62.

[9] HARRIS, D., MCCABE, B.P.M. and LEYBOURNE, S.J. (2002) Stochastic cointegration: estimation and inference. *Journal of Econometrics* 111, 363-84.

[10] HARVEY, A.C. (1991) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

[11] LEYBOURNE, S.J., MCCABE, B.P.M. and TREMAYNE, A.R. (1996) Can economic time series be differenced to stationarity? *Journal of Business and Economic Statistics* 14, 435-46.

[12] MCCABE, B.P.M. and TREMAYNE, A.R. (1995) Testing a time series for difference stationarity. *The Annals of Statistics* 23, 1015-28.

[13] MELARD, G. (1985) Examples of the evolutionary spectrum theory. *Journal of Time Series Analysis* 6, 81-90.

[14] PHILLIPS, P.C.B. (2001) Descriptive econometrics for non-stationary time series with empirical applications. *Journal of Applied Econometrics* 16, 389-413.

[15] PRIESTLEY, M.B. (1981) *Spectral Analysis and Time Series*, Academic Press, London, UK.

[16] PRIESTLEY, M.B. (1988) *Non-linear and Non-stationary Time Series Analysis*, Academic Press, London, UK.

[17] TJØSTHEIM, D. (1976) Spectral generating operators for non-stationary processes', *Advances in Applied Probability*, 8, 83-46.

[18] WOLD, H. (1938) *A Study in the Analysis of Stationary Time Series,* Almqvist and Wiksell, Sweden.