Correlograms for non-stationary autoregressions

By Bent Nielsen
Department of Economics, University of Oxford.
bent.nielsen@nuf.ox.ac.uk
2 April 2003

Analysis of economic time series often involves correlograms and partial correlograms as graphical descriptions of temporal dependence. Two methods are available for computing these statistics: one based on autocorrelations and the other on scaled autocovariances. For a stationary time series the resulting plots are nearly identical. When it comes to economic time series that usually exhibit non-stationary features these methods can lead to very different results. This has two consequences: (i) incorrect inferences can be drawn when confusing these concepts; (ii) a better discrimination between stationary and non-stationarity appears when using autocorrelations rather than autocovariances which are commonly used in econometric software.

Keywords: correlogram, covariogram, non-stationarity.

1 Introduction

The analysis of an economic time series $X_1, \dots, X_T$ often involves correlograms and partial correlograms as graphical descriptions of temporal dependence. Two methods are available for computing these statistics: one based on sample autocorrelations

$$r_u = \frac{\sum_{t=u+1}^{T}(X_t - \overline{X}_{1+u}^{T})(X_{t-u} - \overline{X}_1^{T-u})}{\sqrt{\sum_{t=u+1}^{T}(X_t - \overline{X}_{1+u}^{T})^2 \sum_{t=u+1}^{T}(X_{t-u} - \overline{X}_1^{T-u})^2}}, \tag{1}$$

and one based on scaled sample autocovariances

$$g_u = \frac{\sum_{t=u+1}^{T}(X_t - \overline{X}_1^{T})(X_{t-u} - \overline{X}_1^{T})}{\sum_{t=1}^{T}(X_t - \overline{X}_1^{T})^2}, \tag{2}$$

for $0 \leq u < T$ and where for instance $\overline{X}_1^{T-u}$ is the sample average of $X_1, \dots, X_{T-u}$. To distinguish these methods it is perhaps more appropriate to refer to g_u as yielding scaled covariograms, or simply covariograms.

For a stationary time series the sample variance is of course constant over time rendering the correlograms and covariograms to be nearly identical. When it comes to non-stationary economic time series correlograms and covariograms can be very different. This has two consequences: (i) incorrect inferences can be drawn when following a long established tradition in confusing these concepts; (ii) a better discrimination between stationary and non-stationarity appears when using correlograms rather than the more commonly used covariograms.

Correlograms rather than covariograms were used in early work such as the seminal paper by Yule (1926), the monograph by Wold (1938, p.12), and the analysis of agricultural price series by Kendall (1943). Being concerned with stationary time series Wold (1938, p.12) and Kendall (1945) could then allow themselves to compare sample correlograms with population covariograms. When developing higher order asymptotic theory researchers such as Bartlett (1946), and Quenouille (1947) concentrated on sample covariograms which are more analytically tractable than correlograms, while Anderson (1942) simplified the problem further by restricting the analysis to circular autoregressions. Perhaps as a consequence of this work recent text books such as Brockwell and Davis (1996) and many time series computer programs such as Ox 3, PcGive 10, R 1.3, and RATS 4.3 have adopted covariograms rather than correlograms as the basic descriptive statistics of temporal dependence. Some authors like Hendry (1995) use correlograms and accordingly early versions of PcGive reported correlograms but this was changed in version 10, see Doornik and Hendry (2001, p.259). Apparently this was in response to users complaining the output differed from all other packages', hence the dangers of consensus over verisimilitude!

In the following it is demonstrated how correlograms and covariograms differ for non-stationary time series. As a first illustration some typical economic time series are investigated. Stylised features of these series can be captured either by a first order autoregression or a cumulated random walk. Such series are analysed mathematically.

2 Definitions

While the literature appears to be in agreement on the definition of sample correlograms and covariograms there are several definitions of the corresponding population versions in circulation. These are discussed in the following along with partial correlograms and covariograms.

For the definition of population correlograms the joint distribution of the time series $X_1, \dots, X_T$ has to be specified. When thinking of the time series as a realisation of an infinitely lived process $(X_t)_{t\in\mathbf{Z}}$ this has to be done with some care since the marginal distribution of the vector $(X_1, \dots, X_T)'$ will in general be different from the conditional distribution of $(X_1, \dots, X_T)'$ given the information set at time zero. Having made a choice of distribution population correlograms and covariograms can be defined as

$$\rho_{t,u} = \mathsf{Corr}(X_t, X_{t-u}) = \frac{\mathsf{Cov}(X_t, X_{t-u})}{\sqrt{\mathsf{Var}(X_t)\mathsf{Var}(X_{t-u})}}, \qquad \gamma_{t,u} = \frac{\mathsf{Cov}(X_t, X_{t-u})}{\mathsf{Var}(X_t)}, \tag{3}$$

where Corr, Cov, and Var are the correlation, covariance, and variance defined with respect to the chosen distribution for $X_1, \ldots, X_T$. In general $\rho_{t,u}$ and $\gamma_{t,u}$ will vary both with t and u. The exception is when these are evaluated with respect to a stationary distribution so $\rho_{t,u} = \gamma_{t,u}$ is invariant in t.

The sample partial correlogram, p_u, is defined in terms of the usual sample partial correlation of X_t and X_{t-u} corrected for $X_{t-1}, \ldots, X_{t-u}$ while a general definition of partial covariograms, a_u, is based on the Yule-Walker equations, see Brockwell and Davis (1996, p.57,93). For the sake of the arguments in this paper it suffices to look at the first order statistics $p_1 = r_1$ and $a_1 = g_1$ and the second order statistics

$$p_2 = \frac{r_{2,0,0} - r_{1,0,1} r_{1,1,0}}{\sqrt{(1 - r_{1,0,1}^2)(1 - r_{1,1,0}^2)}}, \qquad a_2 = \frac{g_2 - g_1^2}{1 - g_1^2}, \tag{4}$$

where $r_{u,v,w}$ generalises r_u as the u-th autocorrelation of the time series $X_{1+v}, \ldots, X_{T-w}$ and is given by

$$r_{u,v,w} = \frac{\sum_{t=1+u+v}^{T-w} (X_t - \overline{X}_{1+u+v}^{T-w})(X_{t-u} - \overline{X}_{1+v}^{T-w-u})}{\sqrt{\sum_{t=1+u+v}^{T-w} (X_t - \overline{X}_{1+u+v}^{T-w})^2 \sum_{t=1+u+v}^{T-w} (X_{t-v} - \overline{X}_{1+v}^{T-w-u})^2}}. \tag{5}$$

The population partial autocorrelation is defined as

$$\pi_{t,u} = \mathsf{Corr}(X_t, X_{t-u} | X_{t-1}, \ldots, X_{t-u+1}).$$

This function indicates the order of an autoregression. For an autoregression of order q it holds for $u > q$ that X_t and X_{t-u} are conditionally independent given the intermediate observations $X_{t-1}, \ldots, X_{t-u+1}$ and thus $\pi_{t,u} = 0$. In general it holds $\pi_{t,1} = \rho_{t,1}$ and for a time series with a joint normal distribution the second order partial autocorrelation is

$$\pi_{t,2} = \frac{\rho_{t,2} - \rho_{t,1}\rho_{t-1,1}}{\sqrt{(1 - \rho_{t,1}^2)(1 - \rho_{t-1,1}^2)}}. \tag{6}$$

Population partial covariograms, $\alpha_{t,u}$, are usually defined with respect to a stationary distribution, see Brockwell and Davis (1996, p.43,45,93), while no unique

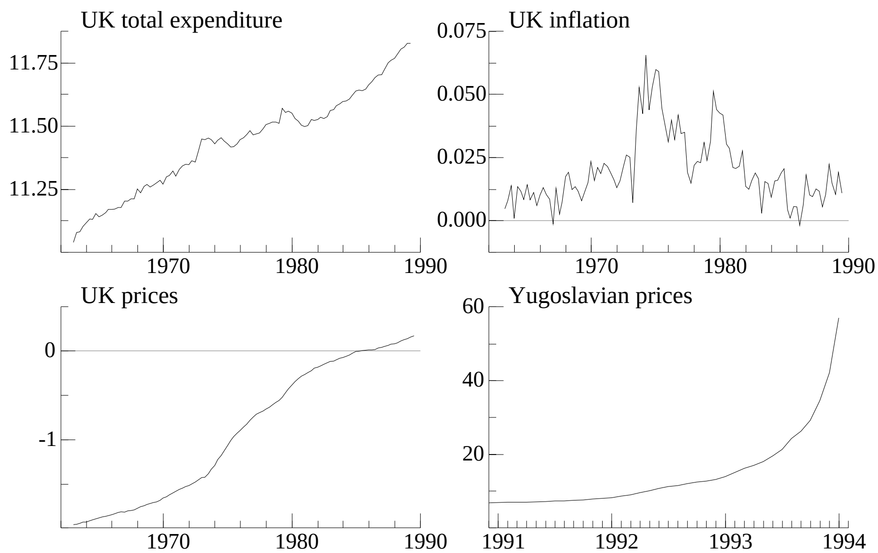

Figure 1: Data series

definition is available for time varying distributions. Here the definitions $\alpha_{t,1} = \gamma_{t,1}$ and

$$\alpha_{t,2} = \frac{\gamma_{t,2} - \gamma_{t,1}^2}{1 - \gamma_{t,1}^2} \tag{7}$$

are used as they give a good match to the sample partial covariograms. The partial covariogram $\alpha_{t,u}$ and the partial correlogram $\pi_{t,u}$ will in general be different. An exception is the case where a stationary normal distributed time series is considered.

3 Correlograms for typical economic time series

To illustrate the different definitions four economic times series are studied. The four series are shown in Figure 1. The first three series are quarterly log prices, inflation measured as differenced log prices and log total expenditure for the UK for the period 1963:1-1989:3. The fourth series is monthly log prices from the Yugoslavian hyperinflation, 1990:12-1994:1. Detailed econometric analysis of the UK series can be found in Doornik, Hendry and Nielsen (1998), which also lists other papers analysing this data set. Petrović and Mladenović (2000) have analysed the Yugoslavian price series.

Rather than discussing detailed congruent models for these data it is perhaps more instructive to discuss stylised models. The total expenditure series and similar series measuring the output have been studied in numerous papers since Nelson and

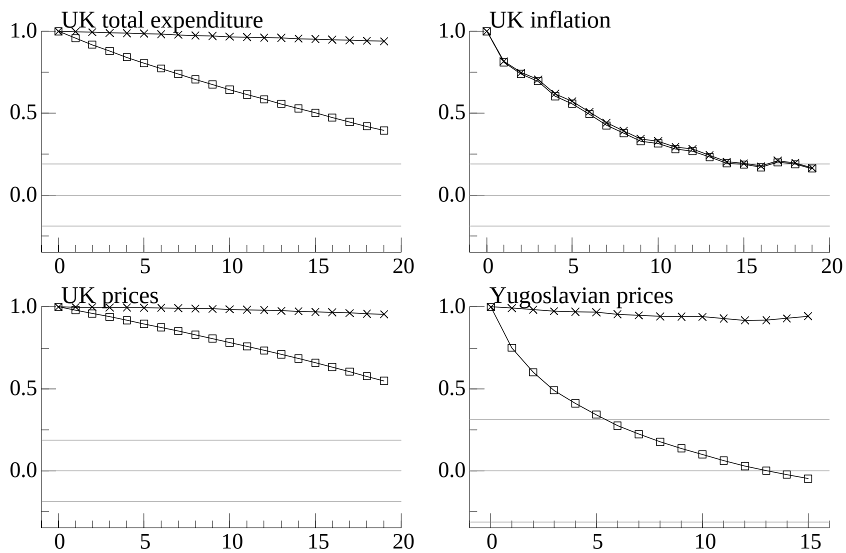

Figure 2: Correlogram, r_u, shown with crosses, covariogram, g_u, shown with boxes. Horizontal lines indicate 95% confidence bands for testing the hypothesis of no serial dependence at a given lag length, see Bartlett (1935), Box and Jenkins (1970, p.35).

Plosser (1982) and are usually described as near $\mathsf{I}(1)$ autoregressions with a linear trend, perhaps with a few dummies to account for events like the fiscal expansions in 1972 and 1979 followed each by an oil crisis, see Doornik, Hendry and Nielsen (1998). In the same way the UK inflation could be described as near $\mathsf{I}(1)$ autoregressive with a constant level and UK prices as near $\mathsf{I}(2)$ autoregressive with a linear trend. Series like the Yugoslavian price series have been studied less in the literature. Following the work of Juselius and Mladenović (2002) it is described as autoregressive with a unit root and an explosive root. To facilitate later discussion these stylised models are summarised as first and second order autoregressions as follows,

$$\begin{aligned} \text{UK expenditure:} && (1-0.9L)X_t &= \varepsilon_t + \text{linear trend.} && (8) \\ \text{UK inflation:} && (1-0.9L)X_t &= \varepsilon_t + \text{constant,} && (9) \\ \text{UK prices:} && (1-0.95L)(1-0.95L)X_t &= \varepsilon_t + \text{linear trend,} && (10) \\ \text{Yugoslavian prices:} && (1-2L)(1-L)X_t &= \varepsilon_t + \text{constant,} && (11) \end{aligned}$$

Sample correlograms and covariograms for the four economic time series are shown in Figure 2. The most striking difference is perhaps for the Yugoslavian prices. The correlogram shows strong persistence while the covariogram is exponentially declining. For UK prices and UK expenditure both methods show strong persistence in that the

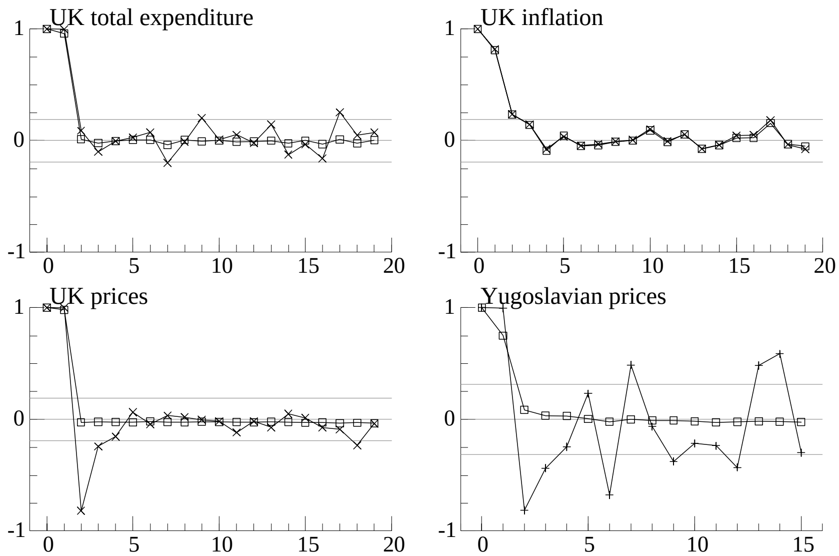

Figure 3: Partial correlogram, p_u, shown with crosses, partial covariogram, a_v, shown with boxes. Horizontal lines are 95% confidence bands for testing the hypothesis of no serial dependence at a given lag length, see Quenouille (1949), Box and Jenkins (1970, p.65).

curves are almost linearly declining although with different slopes. For UK inflation the methods yield more or less the same exponentially declining curve.

Partial correlograms and covariograms are shown in Figure 3. The first impression is that the partial correlogram is much more noisy than the partial covariogram. This is in particular the case for the relatively short Yugoslavian price series. A more important difference is that for the two highly persistent price series the partial covariograms tend to suggest a shorter lag length than the partial correlogram. The stylised models for the two price series given in (11), (10) suggest that at least two lags are needed in agreement with the partial correlogram.

In combination the correlograms and the partial correlograms appear to be able to discriminate the stochastic behaviour of the four time series. As an example the very different stylised models for UK prices and UK total expenditure can be discriminated by the partial correlogram but not very well by the correlogram itself. The two price series are perhaps most difficult to discriminate using the correlograms and partial correlograms and in this instance the covariogram may be useful.

4 Properties of population correlograms

For non-stationary time series population correlograms and covariograms are in general different. This is demonstrated by asymptotic analysis of these functions for a first order autoregression and for a cumulated random walk.

For a first order autoregression

$$X_t = \beta X_{t-1} + \mu + \varepsilon_t, \tag{12}$$

with uncorrelated, standardised innovations ε_t it holds $\mathsf{Cov}(X_t, X_{t-u}) = \beta^u \mathsf{Var}(X_{t-u})$, regardless of the initial value X_0 being random or fixed. The correlogram and covariogram can therefore be expressed entirely in terms of variances,

$$\rho_{t,u} = \beta^u \left\{ \frac{\mathsf{Var}(X_{t-u})}{\mathsf{Var}(X_t)} \right\}^{1/2}, \qquad \gamma_{t,u} = \beta^u \frac{\mathsf{Var}(X_{t-u})}{\mathsf{Var}(X_t)}.$$

When the initial value X_0 is fixed it holds $\mathsf{Var}(X_t) = \sum_{j=0}^{t-1} \beta^{2j}$. Computing this geometric progression and expanding asymptotically for large values of t shows

$$\rho_{t,u} \approx \beta^u, \qquad \gamma_{t,u} \approx \beta^u, \qquad \text{for } |\beta| < 1, \tag{13}$$

$$\rho_{t,u} \approx \beta^u (1 - u/2t), \qquad \gamma_{t,u} \approx \beta^u (1 - u/t), \qquad \text{for } |\beta| = 1, \tag{14}$$

$$\rho_{t,u} \approx 1, \qquad \gamma_{t,u} \approx \beta^{-u}, \qquad \text{for } |\beta| > 1, \tag{15}$$

so for the non-stationary situations $|\beta| \geq 1$ the correlogram declines more slowly than the covariogram.

Consider now a cumulated random walk,

$$Y_t = \sum_{j=1}^{t} \sum_{i=1}^{j} \varepsilon_i, \tag{16}$$

with uncorrelated and standardised innovations. A tedious calculation given in Appendix A shows that

$$\rho_{t,u} \approx 1 - \frac{3u^2}{8t^2}, \qquad \gamma_{t,u} \approx 1 - \frac{3u}{2t}, \tag{17}$$

so once again the correlogram declines more slowly than the covariogram. For this highly non-stationary process it is interesting to study the second order partial autocorrelation and partial scaled autocovariance. For large t these are

$$\pi_{t,2} \approx -1 + \frac{3}{t^2}, \qquad \alpha_{t,2} \approx \frac{-3}{4t}, \tag{18}$$

see Appendix A. The cumulated random walk satisfies a second order autoregressive equation. Unlike the partial autocovariance the partial autocorrelation points to the fact that the process has second order dependence.

5 Properties of sample correlograms

Sample correlograms and covariograms are studied for the first order autoregression (12) and for the cumulated random walk (16). For the latter process also the partial correlograms and covariograms are studied. Proofs are reported in Appendix B.

The first order autoregression (12) with independent and standardised innovations has correlogram and covariogram satisfying

$$r_u \overset{a.s.}{\to} \beta^u, \qquad g_u \overset{a.s.}{\to} \beta^u, \qquad \text{for } |\beta| < 1, \tag{19}$$

$$r_u \overset{a.s.}{\to} 1, \qquad g_u \overset{a.s.}{\to} \beta^{-u}, \qquad \text{for } |\beta| > 1, \tag{20}$$

$$r_u \overset{\mathsf{D}}{\approx} 1 - \frac{u}{2TD}, \qquad g_u \overset{\mathsf{D}}{\approx} 1 + \frac{u(N-A)}{TD}, \qquad \text{for } \beta = 1, \mu = 0, \tag{21}$$

$$r_u \overset{\mathsf{D}}{\approx} 1 - \frac{6u}{\mu^2 T^2}, \qquad g_u \overset{\mathsf{D}}{\approx} 1 - \frac{3u}{T}, \qquad \text{for } \beta = 1, \mu \neq 0. \tag{22}$$

Here the approximations are valid for large values of T and a not too large value of u. In (19), (20) the limiting argument holds almost surely while in (21), (22) convergence in distribution is applied termwise to the asymptotic expansion. The quantities A, D, N are defined in terms of a standard Brownian motion as

$$A = (B_1 - \overline{B})^2, \qquad D = \textstyle\int_0^1 (B_s - \overline{B})^2 ds, \qquad N = \textstyle\int_0^1 (B_s - \overline{B}) dB_s.$$

Following Rothenberg (2002) it holds that N/D and $-1/(2D)$ have the same expectation, so taking expectations term wise in the asymptotic expansions in (21) shows

$$\mathsf{E}T(r_u - 1) \approx -\mathsf{E}(2D)^{-1} > -\mathsf{E}(2D)^{-1} - \mathsf{E}(A/D) \approx \mathsf{E}T(g_u - 1). \tag{23}$$

It is of interest to compare these sample results with the population results reported in §4. In the stationary case there is a perfect match. For the explosive case r_u and $\rho_{t,u}$ both converge to one, while g_u and $\gamma_{t,u}$ have the same general shape. For the random walk case without a linear trend, see (14), (21), the functions are all linear decreasing with a slope of t^{-1} or T^{-1}. Just as $\rho_{t,u} > \gamma_{t,u}$ it holds that on average r_u is larger than g_u in the sense of (23). Turning to the case of a random walk with a linear trend the sample results (22) do not match the population results (14) exactly. The sample correlogram shows much stronger persistence than the population correlogram as one could perhaps expect for a series with a linear trend.

These results facilitate interpretation of the correlograms and covariograms for UK expenditure and inflation given in Figure 2. The stylised model for the inflation series (9) is a near-integrated autoregression with a constant level, so a combination of (19) and (21) is observed. The autoregressive coefficient is so far from unity that the difference (23) between correlogram and covariogram can hardly be seen. For the expenditure series the combination of near-integratedness and a linear trend is so dominating a feature that (22) gives a good guidance to interpreting the plot.

A frequently used diagnostic test for autocorrelation is based on the test statistic suggested by Box and Pierce (1970). In its simplest form the test statistic is $Q = Tg_1^2$ which is asymptotically $\chi^2(1)$ if the time series X_t is a sequence of independent mean zero normal variable with constant variance. The above results for the first order autoregression show that if the alternative is formulated as first order autoregressive dependence then the test based on Q is biased in that the power decreases towards zero as $|\beta| \to \infty$. Using the test statistic Tr_1^2 will give a more reliable test that has power close to one for large values of $|\beta|$ as well as having the same properties as the test based on Q for small values of $|\beta|$.

Turning to the cumulated random walk (16) with independent and standardised innovations the sample correlogram and covariograms have properties matching the sample version in that

$$r_u = 1 - \mathrm{O}_{\mathsf{P}}(T^{-2}), \qquad g_u = 1 - \mathrm{O}_{\mathsf{P}}(T^{-1}). \tag{24}$$

This highly persistent series satisfies a second order regression and provides an example where the sample partial covariogram is not useful in determining the autoregressive order. Looking at the sample partial autocorrelation and autocovariance of second order it holds

$$p_2 = -1 + \mathrm{o}_{\mathsf{P}}(1), \qquad a_2 = \mathrm{o}_{\mathsf{P}}(1), \tag{25}$$

while the time series Y_t satisfies a second order regression. This type of discrepancy was noted in a simulation study by Paulsen and Tjøstheim (1985, Table 4). It matches the population properties found in (18) as well as the actual realisations for both the UK and the Yugoslavian prices seen in Figure 3.

For a stationary autoregression of order q it holds that both autocorrelation and autocovariance vanish for orders higher than q so inferences about the lag length can be drawn from both approaches. As an example Hannan and Quinn (1979) base their information criterion on the partial covariogram. The analysis of Nielsen (2001) shows that inferences based on the partial correlogram are actually valid regardless of whether the autoregression is stationary or not. While those arguments are based on asymptotic analysis the simulation study by Paulsen and Tjøstheim (1985) shows that for finite samples of strongly autocorrelated but stationary autoregressions it is preferable to draw inferences from the partial correlogram. This recommendation is followed by software like PcGive 10 and RATS 4.

6 Conclusions

It has been shown for non-stationary autoregressive time series that incorrect inferences can be drawn when confusing the concepts of correlograms and covariograms. Choosing between the two, correlograms tend to give a better discrimination between stationarity and non-stationarity than covariograms.

7 Acknowledgments

The numerical results were generated using Ox (see Doornik, 1999). Discussions with C. Bos, D.R. Cox, D.F. Hendry and N. Shephard are gratefully acknowledged.

8 References

Anderson, R.L. (1942) Distribution of the serial coefficient. *Annals of Mathematical Statistics*, 13, 1-13.

Bartlett, M.S. (1935) Some aspects of the time-correlation problem in regard to tests of significance. *Journal of the Royal Statistical Society*, 98, 536-543.

Bartlett, M.S. (1946) On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, 8, 27-41.

Box, G.E.P., and Jenkins, G.M. (1970) *Time series analysis: forecasting and control.* San Francisco: Holden-Day.

Box, G.E.P., and Pierce, D.A. (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65, 1509-1526.

Brockwell, P.J., and Davis, R.A. (1996) *Introduction to time series and forecasting.* New York: Springer.

Doornik, J.A. (1999) *Object-oriented matrix programming using Ox*, 3rd ed. London: Timberlake Consultants Press.

Doornik, J.A. and Hendry, D.F. (2001) *Empirical econometric modelling using PcGive 10, vol. 1.* London: Timberlake Consultants Press.

Doornik, J.A., Hendry, D.F. and Nielsen, B. (1998) Inference in cointegrating models: UK M1 revisited. *Journal of Economic Surveys*, 12, 533-572.

Gradshteyn, I. S., and Ryzhik, I. M. (1965) *Tables of integrals, series and products.* New York: Academic Press.

Hannan, E.J., and Quinn, B.G. (1979) The determination of the order of an autoregression. *Journal of the Royal Statistical Society B*, 41, 190-195.

Hendry, D.F. (1995) *Dynamic econometrics.* Oxford: Oxford University Press.

Juselius, K. and Mladenović, Z. (2002) High inflation, hyperinflation and explosive roots. The case of Yugoslavia. Copenhagen: http://www.econ.ku.dk/okokj.

Kendall, M.G. (1943) Oscillatory movements in English agriculture (with discussion). *Journal of the Royal Statistical Society*, 106, 91-124.

Kendall, M.G. (1945) On the analysis of oscillatory time series (with discussion). *Journal of the Royal Statistical Society*, 108, 93-141.

Lai, T.L., and Wei, C.Z. (1983) Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters, *Journal of Multivariate Analysis*, 13, 1-23.

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

Nielsen, B. (2001) Weak consistency of criterions for order determination in a general vector autoregression. Oxford: www.nuff.ox.ac.uk/Users/Nielsen.

Paulsen, J., and Tjøstheim, D. (1985) On the estimation of residual variance and order in autoregressive time series. *Journal of the Royal Statistical Society B*, 47, 216-228.

Petrović, P., and Mladenović, Z. (2000) Money demand and exchange rate determination under hyperinflation: Conceptual issues and evidence from Yugoslavia. *Journal of Money, Credit, and Banking* 32, 785-806.

Phillips, P.C.B., and Solo, V. (1992) Asymptotics for linear processes. *Annals of Statistics,* 20, 971-1001.

Quenouille, M.H. (1947) A large sample test for the goodness of fit of autoregressive schemes. *Journal of the Royal Statistical Society*, 110, 123-129.

Quenouille, M.H. (1949) Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society B*, 11, 68-84.

Rothenberg, T.J. (2002) Some elementary distribution theory for an autoregression fitted to a random walk. *Journal of Econometrics*, 111, 355-361.

Wold, H. (1938) *A study in the analysis of stationary time series.* Uppsala: Almqvist & Wiksell.

Yule, G.U. (1926) Why do we sometimes get nonsense-correlations between time-series? - A study in sampling and the nature of time-series (with discussion). *Journal of the Royal Statistical Society*, 89, 1-63.

A Derivation of the population results for a cumulated random walk

The population results (17), (18) for the cumulated random walk Y_t given in (16) are found by first deriving the variance and covariance and then expand the appropriate functions of those for large values of t.

The variance and covariance of $Y_t = \sum_{j=1}^t \sum_{i=1}^j \varepsilon_i = \sum_{i=1}^t i\varepsilon_{t-i+1}$ are given by

$$\begin{aligned}
\mathsf{Var}(Y_t) &= \textstyle\sum_{i=1}^t i^2 = \frac{1}{6}t(t+1)(2t+1), & (26)\\
\mathsf{Cov}(Y_t, Y_{t-u}) &= \mathsf{Var}(Y_{t-u}) + u\mathsf{Cov}(\textstyle\sum_{i=1}^{t-u}\varepsilon_i, Y_{t-u}) \\
&= \mathsf{Var}(Y_{t-u}) + u\textstyle\sum_{i=1}^{t-v} i = \frac{1}{6}(t-u)(t-u+1)(2t+u+1), & (27)
\end{aligned}$$

where Gradshteyn and Ryzhik (1965, 0.121.2) has been used.

Combine the expressions (26), (27) in the ratios

$$\begin{aligned}
\gamma_{t,u} = \frac{\mathsf{Cov}(Y_t, Y_{t-u})}{\mathsf{Var}(Y_t)} &= \left(1-\frac{u}{t}\right)\left(1-\frac{u}{t+1}\right)\left(1+\frac{u}{2t+1}\right), \\
\frac{\mathsf{Var}(Y_{t-u})}{\mathsf{Var}(Y_t)} &= \left(1-\frac{u}{t}\right)\left(1-\frac{u}{t+1}\right)\left(1-\frac{2u}{2t+1}\right).
\end{aligned}$$

Taylor expanding the first expression for large values of t and fixed values of u gives

$$\gamma_{t,u} = \frac{\mathsf{Cov}(Y_t, Y_{t-u})}{\mathsf{Var}(Y_t)} = 1 - \frac{3u}{2t} + \frac{3u}{4t^2} + \frac{4u^3-7u}{8t^3} - \frac{3u(4u^2-5)}{16t^4} + \mathrm{o}(t^{-5}), \qquad (28)$$

which is the formula for $\gamma_{t,u}$ in (17), while the second expression expands as

$$\frac{\mathsf{Var}(Y_{t-u})}{\mathsf{Var}(Y_t)} = 1 - \frac{3u}{t} + \frac{3u(2u+1)}{2t^2} - \frac{4u^3+12u^2+5u}{4t^3} + \frac{3u(4u^2+8u+3)}{8t^4} + \mathrm{o}(t^{-5}).$$

Applying a Taylor expansion of $(1+x)^{-1/2}$ to the latter expression gives

$$\begin{aligned}
\sqrt{\frac{\mathsf{Var}(Y_t)}{\mathsf{Var}(Y_{t-u})}} &= 1 + \frac{3u}{2t} + \frac{3u(5u-2)}{8t^2} + \frac{5u(7u^2-6u+2)}{16t^3} \\
&\quad + \frac{3u(105u^3-140u^2+92u-24)}{128t^4} + \mathrm{o}(t^{-5}). & (29)
\end{aligned}$$

Multiplying (28) and (29) gives the expression for $\rho_{t,u}$ in (17) which is

$$\rho_{t,u} = 1 - \frac{3u^2}{8t^2} - \frac{u(u-1)(u-2)}{8t^3} + \frac{3u(-3u^3+8u^2-28u+16)}{128t^4} + \mathrm{o}(t^{-5}). \qquad (30)$$

Inserting the expression for $\gamma_{t,u}$ and $\rho_{t,u}$ in (28) and (30) into the definition (7) for $\alpha_{t,2}$ and (6) for $\pi_{t,2}$, respectively, yields the desired expressions (18).

B Derivations of the sample results

At first some general formulas for correlograms and covariograms are derived. Subsequently the particular time series are studied.

The generalised autocorrelation (5) can be written as

$$r_{u,v,w} = \frac{S_{11} + S_{10}}{\sqrt{S_{11}(S_{11} + 2S_{10} + S_{00})}} \tag{31}$$

where $S_{ij} = \sum_{t=1+u+v}^{T-w} R_{i,t} R_{j,t}$ and $R_{1,t}$ and $R_{0,t}$ are defined in terms of

$$(X_t - \overline{X}_{1+u+v}^{T-w}) = (X_{t-u} - \overline{X}_{1+v}^{T-w-u}) + (X_t - X_{t-u} - \overline{X}_{1+u+v}^{T-w} + \overline{X}_{1+v}^{T-w-u}) \stackrel{\text{def}}{=} R_{1,t} + R_{0,t}.$$

Provided that $(S_{10}/S_{11})^2 = \mathrm{O}(S_{00}/S_{11}) = \mathrm{o}(1)$ for large T and fixed u, v, w an asymptotic expansion shows

$$r_{u,v,w} \approx 1 - \frac{S_{00}}{2S_{11}} - \left(\frac{S_{10}}{2S_{11}}\right)^2. \tag{32}$$

If in addition it holds $(S_{10}/S_{11})^2 = \mathrm{o}(S_{00}/S_{11})$ then

$$r_{u,v,w} \approx 1 - \frac{S_{00}}{2S_{11}}. \tag{33}$$

To derive expressions for the covariograms g_u is rewritten in the same way as

$$g_u = \frac{\sum_{t=u+1}^{T}(X_t - \overline{X}_1^T)(X_{t-u} - \overline{X}_1^T)}{\sum_{t=1}^{T}(X_t - \overline{X}_1^T)^2} \stackrel{\text{def}}{=} 1 + \frac{S_2 - S_3}{S_1}, \tag{34}$$

where S_1 is the denominator of g_u and

$$S_2 = \sum_{t=v+1}^{T}(X_t - X_{t-u})(X_{t-u} - \overline{X}_1^T), \qquad S_3 = \sum_{t=T-u+1}^{T}(X_t - \overline{X}_1^T)^2.$$

B.1 Proof of sample expression (19) for the asymptotically stationary case

For $|\beta| < 1$ the result is well-known. It can be proved using laws of large numbers for linear processes, see Phillips and Solo (1992, Theorem 3.1, 3.7), that

$$T^{-1}(S_{11}, S_{10}, S_{00}, S_1, S_2, S_3) \stackrel{a.s.}{\to} (1, \beta^u - 1, 2 - 2\beta^u, 1, \beta^u - 1, 0)/(1 - \beta^2).$$

The desired results for r_u and g_u then follow from (31), (34).

B.2 Proof of sample expression (20) for the explosive case

For $|\beta| > 1$ the distribution of the sample expressions are invariant to μ since $X_t - \mu/(1-\beta)$ satisfy an autoregression without intercept but with initial value $\tilde{X}_0 = X_0 - \mu/(1-\beta)$. Lai and Wei (1983, Theorem 2) prove that

$$\beta^{-t} X_t \stackrel{a.s.}{\to} W, \qquad \beta^{-2T} \sum_{t=1}^{T} X_t^2 \stackrel{a.s.}{\to} W^2/(1-\beta^{-2}),$$

where $W = \tilde{X}_0 + \sum_{t=1}^{\infty} \beta^{-t} \varepsilon_t$ is a continuous random variable. Moreover it holds

$$\beta^{-T} \sum_{t=1}^{T} X_t \stackrel{a.s.}{\to} W/(1-\beta^{-1}), \qquad \beta^{-T} \sum_{t=1}^{T} X_t \varepsilon_{t+j} \stackrel{a.s.}{=} \mathrm{o}(T^{1/2}), \qquad \text{for } j > 0.$$

where the first result is proved as that of $\sum_{t=1}^{T} X_t^2$ and the latter result follows as in Lai and Wei (1983, Equation 4.18). The desired results for r_u and g_u follow from (31) and (34) noting that $X_t - X_{t-u} = \sum_{j=0}^{u-1} \beta^j \varepsilon_{t-j} + (\beta^u - 1) X_{t-u}$ and

$$\begin{aligned} \beta^{-2(T-u-v-w)}(S_{11}, S_{10}, S_{00}) &\stackrel{a.s.}{\to} \{1, (\beta^u - 1), (\beta^u - 1)^2\} W^2/(1-\beta^{-2}), \\ \beta^{-2T}(S_1, S_2, S_3) &\stackrel{a.s.}{\to} \{1, (\beta^u - 1)\beta^{-2u}, 1 - \beta^{-2u}\} W^2/(1-\beta^{-2}). \end{aligned}$$

B.3 Proof of sample expression (21) for random walk with constant level

For $\beta = 1$ and $\mu = 0$ then $\Delta X_t = \varepsilon_t$ and the distributions of r_u and g_u are invariant to X_0, which can then be chosen as $X_0 = 0$. It then holds $X_t - X_{t-u} = \sum_{j=1}^{u} \varepsilon_{t-u+j}$ and $X_{t-u} = \sum_{j=1}^{t-u} \varepsilon_j$. The Law of Large Numbers and Donsker's invariance principle combined with the Continuous Mapping Theorem, see Billingsley (1968), then imply

$$T^{-1}(T^{-1} S_{11}, S_{10}, S_{00}, T^{-1} S_1, S_2, S_3) \stackrel{\mathsf{D}}{\to} (D, uN, u, D, uN, uA).$$

The desired results for r_u and g_u follow by inserting these expressions in (33), (34).

B.4 Proof of sample expression (22) for random walk with linear trend

For $\beta = 1$ and $\mu \neq 0$ then $\Delta X_t = \mu + \varepsilon_t$ and the distributions of the sample expressions are invariant to X_0, which can then be chosen as $X_0 = 0$. It then holds $X_t - X_{t-u} = \mu u + \sum_{j=1}^{u} \varepsilon_{t-u+j}$ and $X_{t-u} = \mu(t-u) + \sum_{j=1}^{t-u} \varepsilon_j$. Noting that a linear trend dominates a random walk the Law of Large Numbers and Donsker's invariance principle combined with the Continuous Mapping Theorem, see Billingsley (1968), then imply

$$T^{-3} S_{11} \stackrel{\mathsf{P}}{\to} \mu^2/12, \qquad S_{10} = \mathrm{O}_{\mathsf{P}}(T^{3/2}), \qquad T^{-1} S_{00} \stackrel{a.s.}{\to} u,$$

$$T^{-3} S_1 \stackrel{\mathsf{P}}{\to} \mu^2/12, \qquad S_2 = \mathrm{O}_{\mathsf{P}}(T^{3/2}), \qquad T^{-2} S_3 \stackrel{\mathsf{P}}{\to} \mu^2 u/4.$$

The desired results for r_u and g_u follow by inserting these expressions in (33), (34).

B.5 Proof of sample expressions (24), (25) for I2 proces

For the cumulated random walk Y_t given by (16) it holds

$$Y_t - Y_{t-u} = u \sum_{i=1}^{t-u} \varepsilon_i + \sum_{j=t-u+1}^{t} \sum_{i=t-u+1}^{j} \varepsilon_i.$$

Let $C_s = \int_0^s B_r dr$ denote the integrated Brownian motion, $\overline{C} = \int_0^1 C_s ds$ and define

$$I_{CC} = \int_0^1 (C_s - \overline{C})^2 ds, \quad I_{CB} = \int_0^1 (C_s - \overline{C})(B_s - \overline{B}) ds, \quad I_{BB} = \int_0^1 (B_s - \overline{B})^2 ds.$$

It then follows from Donsker's invariance principle combined with the Continuous Mapping Theorem, see Billingsley (1968), that

$$T^{-4} S_{11} \overset{\mathsf{D}}{\to} I_{CC}, \qquad T^{-3} S_{10} \overset{\mathsf{D}}{\to} u I_{CB} \qquad T^{-2} S_{00} \overset{\mathsf{D}}{\to} u^2 I_{BB},$$

$$T^{-4} S_1 \overset{\mathsf{D}}{\to} I_{CC}, \qquad T^{-3} S_2 \overset{\mathsf{D}}{\to} u I_{CB}, \qquad T^{-3} S_3 \overset{\mathsf{D}}{\to} u(C_1 - \overline{C})^2.$$

Inserting these expressions in (32), (34) gives the results

$$r_{u,v,w} = 1 - \frac{u^2}{T^2} \left\{ \frac{I_{BB}}{2 I_{CC}} - \left(\frac{I_{CB}}{I_{CC}} \right)^2 \right\} + \mathrm{O}_{\mathsf{P}}(T^{-3}), \quad g_u = 1 - \frac{u I_{CB}}{T I_{CC}} + \mathrm{O}_{\mathsf{P}}(T^{-2}),$$

which are expansions of higher order than what is reported in (24). Inserting these expansions in the definitions of p_2 and a_2 in (4) then gives (25).