LACEA 2001
MONTEVIDEO URUGUAY

Pedro H. Albuquerque

A Simple Nonparametric Long-Run Correlation Estimator

A Simple Nonparametric Long-Run Correlation Estimator

Pedro H. Albuquerque

University of Wisconsin – Madison
Department of Studies and Research – Central Bank of Brazil

Banco Central do Brasil – DEPEP
SBS Q. 3 Bl. A 9º Andar
70074-900 Brasilia DF Brazil
pedrohalbuquerque@yahoo.com
http://go.to/pedroalbuquerque

May 31, 2001

Preliminary

JEL Classification: C14, C22
Keywords: Long-Run Correlation, Nonparametric Estimation, Coherency, Bartlett Kernel, Optimal Lag Selection

Abstract

A simple consistent nonparametric estimator of the long-run correlation between two series is proposed, based on the estimation of the bivariate *k*-lag difference correlation. It is shown that the estimator is asymptotically equivalent to the Bartlett kernel spectral estimator of the complex coherency at frequency zero. The asymptotic distribution is derived, with a test for the absence of long-run correlation. An optimal lag-selection criterion is also presented. Monte Carlo experiments are used to evaluate the proposed estimator.

I am grateful to F. Araújo, S. Durlauf, J. Faria, B. Hansen, M. Horvath, Y. Kitamura, B. Tabak and particularly K. West for many helpful comments and suggestions. I am nevertheless solely responsible for remaining errors.

1 Introduction

This paper proposes a consistent nonparametric estimator for the long-run correlation between two variables. It starts by presenting the concept of long-run correlation between two series. A nonparametric and consistent long-run correlation estimator is considered.

It is shown that an estimator based on the series k-lag difference correlation is asymptotically equivalent to the Bartlett kernel spectral estimator of the complex coherency at frequency zero. The asymptotic distribution is derived, with a test for the absence of long-run correlation.

An optimal lag-selection criterion is also presented. Monte Carlo experiments are used to evaluate the proposed estimator.

2 Long-Run Correlation

The concept of long-run correlation is not new in economics.[1] It can be defined using the complex coherency function from spectral analysis. Granger and Weiss (1983), for example, noted that two $I(1)$ series are cointegrated if and only if the differenced series have coherency equal to one at frequency zero, meaning that their squared long-run correlation is equal to one.[2] Here long-run correlation is used to represent a measure of long-run relationship between two series. This paper will consider only the bivariate case.

Priestley (1981) proposes that if we consider the relationship between the stationary series Δp_t and Δu_t then the coherency at frequency w can be interpreted as the correlation between the random coefficients of the spectral components of Δp_t and Δu_t at frequency w. It measures the degree of linear

---

[1] See for example Granger and Rees (1968), Granger and Engle (1983), McCallum (1984), Sargent (1987), King and Watson (1994), Weber (1994), and King and Watson (1997).
[2] See also Engle and Granger (1987, pg. 254).

dependency between Δp_t and Δu_t for an arbitrarily small frequency interval in the neighborhood of w. The definition of complex coherency is

$$C(w) = \frac{s_{\Delta p \Delta u}(w)}{\sqrt{s_{\Delta p \Delta p}(w) s_{\Delta u \Delta u}(w)}}, \ |C(w)| \le 1,$$

where

$$s_{\Delta p \Delta p}(w) = \frac{1}{2p} \sum_{n=-\infty}^{\infty} g_n(\Delta p_t, \Delta p_t) e^{-iwn},$$

$$s_{\Delta u \Delta u}(w) = \frac{1}{2p} \sum_{n=-\infty}^{\infty} g_n(\Delta u_t, \Delta u_t) e^{-iwn},$$

$$s_{\Delta p \Delta u}(w) = \frac{1}{2p} \sum_{n=-\infty}^{\infty} g_n(\Delta p_t, \Delta u_t) e^{-iwn},$$

$$g_n(\Delta p_t, \Delta p_t) = E[(\Delta p_t - m_{\Delta p})(\Delta p_{t-n} - m_{\Delta p})],$$

$$g_n(\Delta u_t, \Delta u_t) = E[(\Delta u_t - m_{\Delta u})(\Delta u_{t-n} - m_{\Delta u})],$$

$$g_n(\Delta p_t, \Delta u_t) = E[(\Delta p_t - m_{\Delta p})(\Delta u_{t-n} - m_{\Delta u})], \ m_{\Delta p} = E[\Delta p_t], \ m_{\Delta u} = E[\Delta u_t],$$

$s_{\Delta p \Delta p}(w), s_{\Delta u \Delta u}(w)$ are the spectra and $s_{\Delta p \Delta u}(w)$ is the cross-spectrum (the Fourier transform of the autocovariances or the cross-covariances) and g_n is the autocovariance or the cross-covariance of Δp_t and Δu_t.[3]

The time-domain concept of *long-run correlation* is equivalent to the frequency-domain concept of *complex coherency at frequency zero*, as defined below:

Definition 1: *given two $I(0)$ processes Δp_t and Δu_t, define long-run correlation as*

$$l \equiv \frac{s_{\Delta p \Delta u}}{\sqrt{s_{\Delta p \Delta p} s_{\Delta u \Delta u}}}, \ -1 \le l \le 1, \tag{2.1}$$

where

---

[3] References for this topic can be found in Koopmans (1974), Fuller (1976), Priestley (1981), Granger and Watson (1984), Sargent (1987), Nerlove and Diebold (1990), Brockwell and Davis (1991), Reinsel (1993), and Hamilton (1994).

$$s_{\Delta p\Delta u} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} g_n(\Delta p_t, \Delta u_t),\ s_{\Delta p\Delta p} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} g_n(\Delta p_t, \Delta p_t),\ s_{\Delta u\Delta u} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} g_n(\Delta u_t, \Delta u_t).$$

3 Nonparametric Estimation of Long-Run Correlation

This section presents a nonparametric estimator for the long-run correlation between two variables. The approach here is similar to the one used in Cochrane (1988), and Cochrane and Sbordone (1988). Cochrane developed a nonparametric statistic for unit root processes called the variance ratio, based on the *k*-lag difference variance of a series.[4] Here a similar concept is applied to the long-run correlation between two series.

3.1 Estimator

Consider two $I(1)$ series p_t and u_t having long-run correlation equal to l. Their first-difference processes are assumed to have summable covariances and autocovariances. Given a sample of size T+1, $0 \le t \le T$, an analog estimator for the long-run correlation is a kernel estimator

$$\hat{l}_k = \frac{\hat{s}_{\Delta p\Delta u}(k)}{\sqrt{\hat{s}_{\Delta p\Delta p}(k)\hat{s}_{\Delta u\Delta u}(k)}}, \tag{3.1}$$

where

$$\hat{s}_{\Delta p\Delta u}(k) = \frac{1}{2p}\sum_{n=-T+1}^{T-1} k(n,k)\hat{g}_n(\Delta p_t, \Delta u_t),$$

$$\hat{s}_{\Delta p\Delta p}(k) = \frac{1}{2p}\sum_{n=-T+1}^{T-1} k(n,k)\hat{g}_n(\Delta p_t, \Delta p_t),$$

$$\hat{s}_{\Delta u\Delta u}(k) = \frac{1}{2p}\sum_{n=-T+1}^{T-1} k(n,k)\hat{g}_n(\Delta u_t, \Delta u_t),$$

[4] See also Hamilton (1994, pg. 531).

$$\hat{g}_n(\Delta p_t, \Delta u_t) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T+n}(\Delta p_t - \overline{m}_{\Delta p})(\Delta u_{t-n} - \overline{m}_{\Delta u}), & n < 0 \\ \frac{1}{T}\sum_{t=1}^{T-n}(\Delta p_{t+n} - \overline{m}_{\Delta p})(\Delta u_t - \overline{m}_{\Delta u}), & n \geq 0, \end{cases}$$

$$\hat{g}_n(\Delta p_t, \Delta p_t) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T+n}(\Delta p_t - \overline{m}_{\Delta p})(\Delta p_{t-n} - \overline{m}_{\Delta p}), & n < 0 \\ \frac{1}{T}\sum_{t=1}^{T-n}(\Delta p_{t+n} - \overline{m}_{\Delta p})(\Delta p_t - \overline{m}_{\Delta p}), & n \geq 0, \end{cases}$$

$$\hat{g}_n(\Delta u_t, \Delta u_t) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T+n}(\Delta u_t - \overline{m}_{\Delta u})(\Delta u_{t-n} - \overline{m}_{\Delta u}), & n < 0 \\ \frac{1}{T}\sum_{t=1}^{T-n}(\Delta u_{t+n} - \overline{m}_{\Delta u})(\Delta u_t - \overline{m}_{\Delta u}), & n \geq 0, \end{cases}$$

$$\overline{m}_{\Delta p} = \frac{1}{T}\sum_{t=1}^{T}\Delta p_t \, , \; \overline{m}_{\Delta u} = \frac{1}{T}\sum_{t=1}^{T}\Delta u_t \, ,$$

and $k(n,k)$ is a kernel with bandwidth k, $1 \leq k \leq T-1$.[5] Different kernels can be used here, each one with its pros and cons. Here the Bartlett kernel is chosen because it will lead to a simple time-domain representation of the estimator. The Bartlett kernel is defined as

$$k(n,k) = \begin{cases} 1 - |n|/k, & |n| < k \\ 0, & |n| \geq k, \end{cases}$$

and thereafter

$$\hat{s}_{\Delta p \Delta u}(k) = \frac{1}{2p}\sum_{n=-k+1}^{k-1}\left[\left(1 - \frac{|n|}{k}\right)\hat{g}_n(\Delta p_t, \Delta u_t)\right], \tag{3.2}$$

meaning that, for each choice of the kernel bandwidth k, we will have a different estimator.

Consider now the expression for the k-lag difference covariance estimator divided by k:

---

[5] See Priestley (1981, pg. 704). Note that possible deterministic trends in the levels are extracted through the subtraction of the mean values from the first differences.

$$\frac{\hat{s}_{(1-L^k)p\,(1-L^k)u}}{k} = \frac{1}{k}\sum_{t=k}^{T}\frac{\left[\left(1-L^k\right)p_t - k\bar{m}_{\Delta p}\right]\left[\left(1-L^k\right)u_t - k\bar{m}_{\Delta u}\right]}{T-k},\ \left(1-L^k\right)u_t = u_t - u_{t-k}\,.$$ [6]

This estimator is asymptotically equivalent to the summation in equation (3.2) (the proof is presented in Appendix 1), meaning that:

$$\lim_{T\to\infty,\,k/T\to 0}\frac{\hat{s}_{(1-L^k)p\,(1-L^k)u}}{k} = \lim_{T\to\infty,\,k/T\to 0} 2p\cdot\hat{s}_{\Delta p\Delta u}(k). \tag{3.3}$$

Equation (3.3) shows that the *k*-lag difference correlation estimator can be used to estimate the long-run correlation instead of equation (3.1):

$$\hat{l}_k = \frac{\hat{s}_{(1-L^k)p\,(1-L^k)u}}{\sqrt{\hat{s}_{(1-L^k)p\,(1-L^k)p}\,\hat{s}_{(1-L^k)u\,(1-L^k)u}}}\,. \tag{3.4}$$

The *k*-lag difference estimator has some convenient properties. First, its calculation is simple. Moreover, it has a straightforward time-domain interpretation, measuring the correlation level of the series at *k*-period horizons.

Since the *k*-lag difference estimator and the Bartlett kernel estimator are asymptotically equivalent, the asymptotic properties of the latter can be used as an approximation for the asymptotic properties of the former.

3.2 Consistency

As proven in Newey and West (1987), under certain regularity and mixing conditions the Bartlett kernel estimator of the spectral matrix at frequency zero will be consistent if the bandwidth as a function of the sample size has the properties:

$$\lim_{T\to\infty} k(T) = \infty \text{ and } \lim_{T\to\infty}\frac{k(T)}{T^{1/4}} = 0\,.$$

---

[6] Here, just like in Cochrane (1988, pg. 907), a heuristic finite sample correction that consists of substituting the average of the series' *k*-lag differences by *k* times the average of the series' first differences is introduced.

Moreover, if the process driving the variables can be represented by an infinite-order moving average that has summable coefficients and i.i.d. innovations with finite fourth moments, then it is enough that:

$$\lim_{T\to\infty} k(T) = \infty \text{ and } \lim_{T\to\infty} \frac{k(T)}{T^{1/2}} = 0 .$$

Given the asymptotic equivalence between the Bartlett kernel estimator and the k-lag difference estimator, and using the Slutsky theorem, the consistency properties above are also valid when applied to the k-lag difference estimator of the long-run correlation.

3.3 Asymptotic Distribution

To find the asymptotic distribution of $\hat{l}_k$ under consistency ($k \to \infty$ and $k/T^{1/4} \to 0$), consider first the asymptotic distributions of the Bartlett kernel spectra and cross-spectrum estimators at frequency zero:[7]

$$\sqrt{\frac{T}{k}}\begin{bmatrix} \hat{s}_{\Delta p\Delta p}(k) - s_{\Delta p\Delta p} \\ \hat{s}_{\Delta u\Delta u}(k) - s_{\Delta u\Delta u} \\ \hat{s}_{\Delta p\Delta u}(k) - s_{\Delta p\Delta u} \end{bmatrix} \sim N\left(\mathbf{0}, \frac{2}{3}\begin{bmatrix} 2s_{\Delta p\Delta p}^2 & 2s_{\Delta p\Delta u}^2 & 2s_{\Delta p\Delta p}s_{\Delta p\Delta u} \\ 2s_{\Delta p\Delta u}^2 & 2s_{\Delta u\Delta u}^2 & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u} \\ 2s_{\Delta p\Delta p}s_{\Delta p\Delta u} & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u} & s_{\Delta p\Delta p}s_{\Delta u\Delta u} + s_{\Delta p\Delta u}^2 \end{bmatrix}\right).$$

Now, using the delta method we have:

$$\sqrt{\frac{T}{k}}\left(\hat{l}_k - l\right) \sim N\left(0, \frac{2}{3}\left(1 - l^2\right)^2\right). \tag{3.5}$$

The algebra is presented in Appendix 2. The variance of $\hat{l}_k$ can be estimated using the point estimator of l:

$$\operatorname{var}\left(\hat{l}_k\right) = \frac{2}{3}\frac{k}{T}\left(1 - \hat{l}_k^2\right)^2 . \tag{3.6}$$

[7] See Hannan (1970), Anderson (1971), Fuller (1976), Priestley (1981), and Brockwell and Davis (1991).

3.4 Test of Zero Long-Run Correlation

Under the null of zero long-run correlation the standardized distribution of $\hat{l}_k$ is given by:

$$\sqrt{\frac{3}{2}\frac{T}{k}}\;\hat{l}_k \sim N(0,1), \tag{3.7}$$

meaning that the null hypothesis can be easily tested using the statistic above.

3.5 Lag Selection

The optimal lag selection involves a trade-off between the estimator variance and bias; see Priestley (1981, pg. 517). A larger lag size k (bandwidth) results in higher variance and lower bias, while a smaller lag size k results in lower variance and higher bias. The lag-selection criterion presented below is based on the minimization of the MSE (mean squared error):

$$k = 1.4422\left(\frac{ly}{1-l^2}\right)^{\frac{2}{3}} T^{\frac{1}{3}}.$$

The proof is presented in Appendix 3. The approach is the same of Newey and West (1994).

3.6 Monte Carlo Evaluation

Appendix 4 provides Monte Carlo results for experiments consisting of two Gaussian random walks with different levels of long-run correlation. Each experiment used 100,000 iterations. Different sample sizes and bandwidths were employed. The process is described by:

$$\Delta u_t = a\Delta p_t + (1-a)e_t,\ \Delta p_t \sim \mathrm{NID}(0,1),\ e_t \sim \mathrm{NID}(0,1),$$

where the value of a determines the value of l. The results are symetric for negative long-run correlations.

Table 1 describes the results for the mean point estimates, given by:

$$\frac{1}{100000}\sum_{n=1}^{100000}\hat{l}_{kn}\ .$$

We see that the estimator has a small and downward finite-sample bias for this simple process. In a more elaborate case, where the short-run correlation differs from the long-run correlation, the bias could be positive or negative.

Table 2 shows the standardized variances of the estimates:

$$\frac{3}{2}\frac{T}{k}\left(1-l^2\right)^{-2}\frac{1}{100000}\sum_{n=1}^{100000}\left(\hat{l}_{kn}-\overline{\hat{l}}_k\right)^2.$$

We see that the asymptotic values calculated through equation (3.6) generally provide a good approximation even for small samples, mainly when the bandwidth is between 5 and 20% of the sample size.

Table 3 presents the size and power of a test for the null of zero long-run correlation based on estimated 95% confidence intervals for the long-run correlation. The probabilities are given by:

$$1-p\left[\hat{l}_{kn}-1.96\sqrt{\frac{2}{3}\frac{k}{T}}\left(1-\hat{l}_{kn}^2\right)<0<\hat{l}_{kn}+1.96\sqrt{\frac{2}{3}\frac{k}{T}}\left(1-\hat{l}_{kn}^2\right)\right].$$

The size of this test is not controlled. It tends anyway to be smaller for bandwidths between 5 and 20% of the sample size. The power of this test, although small in some cases, is generally large enough to justify its use.

Finally, Table 4 shows the power and size of the test presented in subsection 3.4:

$$1-p\left[-1.96\sqrt{\frac{2}{3}\frac{k}{T}}<\hat{l}_{kn}<+1.96\sqrt{\frac{2}{3}\frac{k}{T}}\right].$$

This test has a controlled size of 5%. The Monte Carlo results show that for bandwidths between 5 and 10% the size is generally preserved. The power of this test is small for a bandwidth of 20% of the sample size, and is approximately zero for bandwidths larger than 20%.

The Monte Carlo experiments show that the asymptotic approximations found in this section are generally satisfactory even for small samples, particularly for bandwidths between 5 and 20%.

3.7 Advantages and Disadvantages of the Long-Run Correlation Nonparametric Estimator

A nonparametric method has the advantage of being specification free. If we compare the estimator developed in this paper with parametric techniques (VAR, for example), we see that here we do not need to consider the specification of the short-run dynamics or the choice of the number of lags. On the other hand, when compared with well-specified parametric methods, nonparametric methods tend to be less efficient. It should be noted anyway that in general we do not know the correct specification of a model.

The estimation of the long-run correlation in place of the usual estimation of structural long-run parameters has some advantages and disadvantages, depending on the analysis goals. For example, the point estimate and the statistical significance of a long-run parameter can be important in some contexts, even when the long-run correlation is small.

Estimates of the long-run parameters nevertheless do not provide any specific information about the relative importance of the long-run relation between two series. The long-run correlation estimator however does this. For example, the squared long-run correlation of two cointegrated series is equal to one, indicating that the long-run changes of both series follow a perfect linear relationship.

Additionally, the long-run correlation estimator has the advantage of being a standardized measure of relationship, useful for comparisons between economic units. This is one of the reasons why it is commonly used in place of the long-run covariance estimator. For example, although the latter can be employed to test long-run neutrality propositions for a group of countries, it is not standardized and thereafter its values cannot be easily compared.

Conclusions

A simple nonparametric and consistent estimator of the long-run correlation was presented. It was defined as the correlation between the *k*-lag differences of two series, and it was shown to be asymptotically equivalent to the Bartlett kernel spectral estimator of the complex coherency at frequency zero. The properties of the estimator were defined, including the asymptotic distribution and a simple test for the null of zero long-run correlation. An optimal lag-selection criterion was developed. Monte Carlo experiments showed that the asymptotic properties are a good approximation even in the case of small samples.

Appendix 1

The k-lag difference covariance estimator of p_t and u_t is:

$$\hat{s}_{(1-L^k)p\,(1-L^k)u} = \frac{1}{T-k}\sum_{t=k}^{T}\left[\left(\left(1-L^k\right)p_t - k\overline{m}_{\Delta p}\right)\right]\left[\left(1-L^k\right)u_t - k\overline{m}_{\Delta u}\right],\ \left(1-L^k\right)u_t = u_t - u_{t-k}$$

$$= \frac{1}{T-k}\sum_{t=k}^{T}\left[\left(\sum_{i=1}^{k}\Delta p'_{t-k+i}\right)\left(\sum_{j=1}^{k}\Delta u'_{t-k+j}\right)\right],\quad \Delta u'_t = \Delta u_t - \overline{m}_{\Delta u}$$

$$= \frac{1}{T-k}\sum_{t=k}^{T}\left[\sum_{l=1}^{k-1}\left(\sum_{i=1}^{l}\Delta p'_{t-k+i}\Delta u'_{t-l+i} + \sum_{j=1}^{l}\Delta p'_{t-l+j}\Delta u'_{t-k+j}\right) + \sum_{l=1}^{k}\Delta p'_{t-k+l}\Delta u'_{t-k+l}\right]$$

$$= \frac{1}{T-k}\left(\sum_{t=k}^{T}\Delta p'_{t-k+1}\Delta u'_t + 2\sum_{t=k-1}^{T}\Delta p'_{t-k+2}\Delta u'_t + \ldots + (k-1)\sum_{t=2}^{T}\Delta p'_{t-1}\Delta u'_t\right.$$

$$\left. + k\sum_{t=1}^{T}\Delta p'_t\Delta u'_t + (k-1)\sum_{t=2}^{T}\Delta p'_t\Delta u'_{t-1} + \ldots + 2\sum_{t=k-1}^{T}\Delta p'_t\Delta u'_{t-k+2} + \sum_{t=k}^{T}\Delta p'_t\Delta u'_{t-k+1}\right) + R(T,k),$$

$$R(T,k) = -\frac{1}{T-k}\left[\sum_{i=2}^{k-1}\sum_{j=1}^{i-1}(i-j)\left(\Delta p'_j\Delta u'_{k-i+j} + \Delta p'_{T-k+i-j+1}\Delta u'_{T-j+1}\right.\right.$$

$$\left.\left. + \Delta p'_{k-i+j}\Delta u'_j + \Delta p'_{T-j+1}\Delta u'_{T-k+i-j+1}\right) + \sum_{j=1}^{k-1}(k-j)\left(\Delta p'_j\Delta u'_j + \Delta p'_{T-j+1}\Delta u'_{T-j+1}\right)\right],$$

$$\text{or}\quad \frac{\hat{s}_{(1-L^k)p\,(1-L^k)u}}{k} = \frac{T}{T-k}2p\cdot\hat{s}_{\Delta p\Delta u}(k) + \frac{R(T,k)}{k}. \qquad \text{(A1.1)}$$

Given the assumption of summable autocovariances and covariances, then $\lim_{T\to\infty,k/T\to 0}\left[R(T,k)/k\right] = 0$. Under these limit conditions, equation (A1.1) is asymptotically equivalent to $2p$ times the Bartlett kernel estimator of the cross-spectrum at frequency zero:

$$\lim_{T\to\infty,k/T\to 0}\frac{\hat{s}_{(1-L^k)p\,(1-L^k)u}}{k} = \lim_{T\to\infty,k/T\to 0}2p\cdot\hat{s}_{\Delta p\Delta u}(k) = \lim_{T\to\infty,k/T\to 0}\sum_{n=-k+1}^{k-1}\left[\left(1-\frac{|n|}{k}\right)\hat{g}_n\left(\Delta p_t,\Delta u_t\right)\right].$$

Appendix 2

From Priestley (1981, pg. 699) and Brockwell and Davis (1991, pg. 446) we know that the asymptotic covariance between two spectra or cross-spectra estimators at frequency zero is given by:

$$\operatorname{cov}\left(\hat{s}_{ab}(k),\hat{s}_{cd}(k)\right)=\frac{2}{3}\frac{k}{T}\left(s_{ac}s_{bd}+s_{ad}s_{bc}\right). \tag{A2.1}$$

The long-run correlation and its estimator are defined as:

$$l=\frac{s_{\Delta p\Delta u}}{\sqrt{s_{\Delta p\Delta p}\,s_{\Delta u\Delta u}}},\quad \hat{l}_k=\frac{\hat{s}_{\Delta p\Delta u}(k)}{\sqrt{\hat{s}_{\Delta p\Delta p}(k)\hat{s}_{\Delta u\Delta u}(k)}}. \tag{A2.2}$$

According to (A2.1), the components of (A2.2) have the following standardized asymptotic distribution, valid for $k\to\infty$ and $k/T^{1/4}\to 0$:

$$\sqrt{\frac{T}{k}}\begin{bmatrix}\hat{s}_{\Delta p\Delta p}(k)-s_{\Delta p\Delta p}\\ \hat{s}_{\Delta u\Delta u}(k)-s_{\Delta u\Delta u}\\ \hat{s}_{\Delta p\Delta u}(k)-s_{\Delta p\Delta u}\end{bmatrix}\sim N\left(\mathbf{0},\frac{2}{3}\begin{bmatrix}2s^2_{\Delta p\Delta p} & 2s^2_{\Delta p\Delta u} & 2s_{\Delta p\Delta p}s_{\Delta p\Delta u}\\ 2s^2_{\Delta p\Delta u} & 2s^2_{\Delta u\Delta u} & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u}\\ 2s_{\Delta p\Delta p}s_{\Delta p\Delta u} & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u} & s_{\Delta p\Delta p}s_{\Delta u\Delta u}+s^2_{\Delta p\Delta u}\end{bmatrix}\right).$$

Thereafter, using the delta method we have:

$$\sqrt{\frac{T}{k}}\left(\hat{l}_k-l\right)\sim N\left(0,\mathbf{D}\cdot\frac{2}{3}\begin{bmatrix}2s^2_{\Delta p\Delta p} & 2s^2_{\Delta p\Delta u} & 2s_{\Delta p\Delta p}s_{\Delta p\Delta u}\\ 2s^2_{\Delta p\Delta u} & 2s^2_{\Delta u\Delta u} & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u}\\ 2s_{\Delta p\Delta p}s_{\Delta p\Delta u} & 2s_{\Delta u\Delta u}s_{\Delta p\Delta u} & s_{\Delta p\Delta p}s_{\Delta u\Delta u}+s^2_{\Delta p\Delta u}\end{bmatrix}\cdot\mathbf{D}'\right),$$

where $\mathbf{D}=\begin{bmatrix}\dfrac{\partial l}{\partial s_{\Delta p\Delta p}} & \dfrac{\partial l}{\partial s_{\Delta u\Delta u}} & \dfrac{\partial l}{\partial s_{\Delta p\Delta u}}\end{bmatrix}=l\begin{bmatrix}\dfrac{-1}{2s_{\Delta p\Delta p}} & \dfrac{-1}{2s_{\Delta u\Delta u}} & \dfrac{1}{s_{\Delta p\Delta u}}\end{bmatrix}$, or

$$\sqrt{\frac{T}{k}}\left(\hat{l}_k-l\right)\sim N\left(0,\frac{2}{3}l^2\begin{bmatrix}\dfrac{-1}{2s_{\Delta p\Delta p}} & \dfrac{-1}{2s_{\Delta u\Delta u}} & \dfrac{1}{s_{\Delta p\Delta u}}\end{bmatrix}\begin{bmatrix}s_{\Delta p\Delta p}-\dfrac{s^2_{\Delta p\Delta u}}{s_{\Delta u\Delta u}}\\ s_{\Delta u\Delta u}-\dfrac{s^2_{\Delta p\Delta u}}{s_{\Delta p\Delta p}}\\ -s_{\Delta p\Delta u}+\dfrac{s_{\Delta p\Delta p}s_{\Delta u\Delta u}}{s_{\Delta p\Delta u}}\end{bmatrix}\right)\Rightarrow$$

$$\sqrt{\frac{T}{k}}\left(\hat{l}_k-l\right)\sim N\left(0,\frac{2}{3}l^2\left(\frac{1}{l^2}-2+l^2\right)\right)\Rightarrow\sqrt{\frac{T}{k}}\left(\hat{l}_k-l\right)\sim N\left(0,\frac{2}{3}\left(1-l^2\right)^2\right).$$

Appendix 3

Problem: $\min_k \text{MSE}$,

$$\text{MSE} = \left[\text{bias}\left(\hat{l}_k\right)\right]^2 + \text{var}\left(\hat{l}_k\right). \tag{A3.1}$$

Asymptotically:

$$\text{var}\left(\hat{l}_k\right) = \frac{k}{T}\frac{2}{3}\left(1 - l^2\right)^2. \tag{A3.2}$$

And the bias is given by:

$$\text{bias}\left(\hat{l}_k\right) = E\left[\hat{l}_k - l\right] = E\left[\frac{\hat{s}_{\Delta p \Delta u}}{\sqrt{\hat{s}_{\Delta p \Delta p}\,\hat{s}_{\Delta u \Delta u}}} - \frac{s_{\Delta p \Delta u}}{\sqrt{s_{\Delta p \Delta p}\, s_{\Delta u \Delta u}}}\right].$$

Using Taylor series to linearly expand this expression at coordinates $s_{\Delta p \Delta u}$, $s_{\Delta p \Delta p}$ and $s_{\Delta u \Delta u}$:

$$\text{bias}\left(\hat{l}_k\right) \approx E\left[\frac{1}{\sqrt{s_{\Delta p \Delta p}\, s_{\Delta u \Delta u}}}\left(\hat{s}_{\Delta p \Delta u} - s_{\Delta p \Delta u}\right) - \frac{1}{2}\frac{s_{\Delta p \Delta u}}{s^{3/2}_{\Delta p \Delta p}\, s^{1/2}_{\Delta u \Delta u}}\left(\hat{s}_{\Delta p \Delta p} - s_{\Delta p \Delta p}\right) - \frac{1}{2}\frac{s_{\Delta p \Delta u}}{s^{1/2}_{\Delta p \Delta p}\, s^{3/2}_{\Delta u \Delta u}}\left(\hat{s}_{\Delta u \Delta u} - s_{\Delta u \Delta u}\right)\right],$$

implying that

$$\text{bias}\left(\hat{l}_k\right) \approx E\left[\frac{\hat{s}_{\Delta p \Delta u} - s_{\Delta p \Delta u}}{\sqrt{s_{\Delta p \Delta p}\, s_{\Delta u \Delta u}}} - \frac{1}{2}\, l\left(\frac{\hat{s}_{\Delta p \Delta p} - s_{\Delta p \Delta p}}{s_{\Delta p \Delta p}} + \frac{\hat{s}_{\Delta u \Delta u} - s_{\Delta u \Delta u}}{s_{\Delta u \Delta u}}\right)\right],$$

and since it is known that asymptotically (from Hannan, pg. 283)

$$E\left[\hat{s}_{\Delta p \Delta u} - s_{\Delta p \Delta u}\right] = -\frac{1}{k} s^{(1)}_{\Delta p \Delta u}, \text{ where } s^{(1)}_{\Delta p \Delta u} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} |n| g_n\left(\Delta p_t, \Delta u_t\right),$$

$$E\left[\hat{s}_{\Delta p \Delta p} - s_{\Delta p \Delta p}\right] = -\frac{1}{k} s^{(1)}_{\Delta p \Delta p}, \text{ where } s^{(1)}_{\Delta p \Delta p} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} |n| g_n\left(\Delta p_t, \Delta p_t\right),$$

$$E\left[\hat{s}_{\Delta u \Delta u} - s_{\Delta u \Delta u}\right] = -\frac{1}{k} s^{(1)}_{\Delta u \Delta u}, \text{ where } s^{(1)}_{\Delta u \Delta u} = \frac{1}{2p}\sum_{n=-\infty}^{\infty} |n| g_n\left(\Delta u_t, \Delta u_t\right),$$

then asymptotically:

$$\text{bias}\left(\hat{l}_k\right) = \left[-\frac{1}{k} s^{(1)}_{\Delta p \Delta u}\frac{s_{\Delta p \Delta u}}{s_{\Delta p \Delta u}}\frac{1}{\sqrt{s_{\Delta p \Delta p}\, s_{\Delta u \Delta u}}} + \frac{1}{2}\frac{l}{k}\left(\frac{s^{(1)}_{\Delta p \Delta p}}{s_{\Delta p \Delta p}} + \frac{s^{(1)}_{\Delta u \Delta u}}{s_{\Delta u \Delta u}}\right)\right] \Rightarrow$$

$$\text{bias}\left(\hat{l}_k\right) = -\frac{l}{k}y \quad , y \quad = \left[\frac{s^{(1)}_{\Delta p \Delta u}}{s_{\Delta p \Delta u}} - \frac{1}{2}\left(\frac{s^{(1)}_{\Delta p \Delta p}}{s_{\Delta p \Delta p}} + \frac{s^{(1)}_{\Delta u \Delta u}}{s_{\Delta u \Delta u}}\right)\right]. \qquad \text{(A3.3)}$$

From (A3.1), (A3.2) and (A3.3), asymptotically:

$$\text{MSE} = \frac{1}{k^2}\, l^2 y \quad ^2 + \frac{k}{T}\frac{2}{3}\left(1 - l^2\right)^2 .$$

Minimizing:

$$\frac{\partial\, \text{MSE}}{\partial k} = -2\frac{1}{k^3}\, l^2 y \quad ^2 + \frac{1}{T}\frac{2}{3}\left(1 - l^2\right)^2 = 0$$

gives the following optimal lag:

$$k = 1.4422\left(\frac{ly}{1 - l^2}\right)^{\frac{2}{3}} T^{\frac{1}{3}} .$$

Appendix 4

Table 1 - Monte Carlo - Mean Value of $\hat{l}_k$ (a)

l	T	k/T					
		0.025	0.050	0.100	0.200	0.400	0.800
0.0	80	0.001	0.000	0.000	-0.001	-0.001	0.000
0.0	160	-0.001	0.000	0.000	0.000	0.000	-0.001
0.0	320	0.000	-0.001	-0.001	0.000	0.001	-0.001
0.0	640	0.000	0.000	0.000	0.000	-0.001	0.001
0.0	1280	0.000	0.000	0.001	0.001	0.002	0.001
0.2	80	0.186	0.191	0.193	0.190	0.184	0.175
0.2	160	0.194	0.195	0.194	0.188	0.180	0.178
0.2	320	0.198	0.198	0.196	0.191	0.183	0.178
0.2	640	0.198	0.196	0.193	0.187	0.180	0.175
0.2	1280	0.197	0.195	0.192	0.186	0.178	0.172
0.4	80	0.375	0.383	0.385	0.378	0.366	0.352
0.4	160	0.389	0.392	0.388	0.379	0.364	0.357
0.4	320	0.396	0.395	0.391	0.380	0.365	0.358
0.4	640	0.397	0.394	0.389	0.378	0.363	0.358
0.4	1280	0.397	0.394	0.388	0.377	0.361	0.355
0.6	80	0.568	0.579	0.581	0.574	0.557	0.540
0.6	160	0.588	0.591	0.587	0.576	0.558	0.548
0.6	320	0.595	0.593	0.588	0.574	0.555	0.550
0.6	640	0.596	0.594	0.587	0.574	0.555	0.545
0.6	1280	0.596	0.593	0.586	0.573	0.553	0.541
0.8	80	0.774	0.783	0.785	0.778	0.764	0.746
0.8	160	0.790	0.793	0.790	0.780	0.764	0.754
0.8	320	0.796	0.795	0.789	0.778	0.761	0.755
0.8	640	0.797	0.795	0.790	0.779	0.761	0.753
0.8	1280	0.798	0.795	0.789	0.777	0.759	0.750

(a) Gaussian random walks, 100,000 iterations each experiment;
model: $\Delta u_t = a\Delta p_t + (1-a)e_t$, $\Delta p_t \sim \mathrm{NID}(0,1)$, $e_t \sim \mathrm{NID}(0,1)$;
value of a determines the value of l;
statistic: $\frac{1}{100000}\sum_{n=1}^{100000}\hat{l}_{kn}$.

Table 2 - Monte Carlo - Standardized Variance of $\hat{l}_k$ (a)

l	T	K/T					
		0.025	0.050	0.100	0.200	0.400	0.800
0.0	80	8.906	2.931	1.331	0.950	0.750	0.523
0.0	160	3.707	1.467	1.013	0.924	0.776	0.461
0.0	320	1.730	1.070	0.983	0.958	0.815	0.467
0.0	640	1.144	0.993	0.999	0.992	0.843	0.494
0.0	1280	1.017	0.995	1.015	1.007	0.858	0.514
0.2	80	9.213	3.028	1.372	0.975	0.772	0.543
0.2	160	3.771	1.484	1.031	0.948	0.802	0.477
0.2	320	1.736	1.070	0.985	0.978	0.840	0.481
0.2	640	1.163	1.008	1.009	1.003	0.862	0.510
0.2	1280	1.025	1.012	1.037	1.034	0.885	0.536
0.4	80	9.988	3.207	1.439	1.047	0.850	0.611
0.4	160	3.932	1.537	1.071	1.008	0.879	0.529
0.4	320	1.787	1.096	1.026	1.045	0.919	0.536
0.4	640	1.166	1.018	1.052	1.083	0.957	0.568
0.4	1280	1.022	1.013	1.065	1.096	0.970	0.594
0.6	80	11.593	3.574	1.577	1.172	1.008	0.764
0.6	160	4.201	1.603	1.141	1.128	1.041	0.646
0.6	320	1.820	1.129	1.097	1.183	1.109	0.651
0.6	640	1.191	1.059	1.123	1.223	1.143	0.702
0.6	1280	1.046	1.062	1.143	1.236	1.174	0.744
0.8	80	14.772	4.229	1.842	1.449	1.377	1.168
0.8	160	4.826	1.761	1.270	1.370	1.421	0.945
0.8	320	1.937	1.206	1.226	1.462	1.541	0.958
0.8	640	1.233	1.118	1.255	1.500	1.596	1.024
0.8	1280	1.087	1.129	1.296	1.559	1.659	1.098

(a) Gaussian random walks, 100,000 iterations each experiment;
model: $\Delta u_t = a\,\Delta p_t + (1-a)e_t$, $\Delta p_t \sim \mathrm{NID}(0,1)$, $e_t \sim \mathrm{NID}(0,1)$;
value of a determines the value of l;
statistic: $\frac{3}{2}\frac{T}{k}\left(1-l^2\right)^{-2}\frac{1}{100000}\sum_{n=1}^{100000}\left(\hat{l}_{kn}-\overline{\hat{l}}_k\right)^2$.

Table 3 - Monte Carlo - Confidence Interval Test [(a)]

l	T	k/T					
		0.025	0.050	0.100	0.200	0.400	0.800
				Size			
0.0	80	0.564	0.325	0.173	0.156	0.189	0.216
0.0	160	0.332	0.151	0.110	0.147	0.200	0.169
0.0	320	0.160	0.090	0.105	0.158	0.218	0.173
0.0	640	0.084	0.076	0.108	0.166	0.230	0.192
0.0	1280	0.067	0.078	0.111	0.169	0.239	0.209
				Power			
0.2	80	0.627	0.421	0.264	0.215	0.228	0.244
0.2	160	0.489	0.304	0.210	0.208	0.238	0.198
0.2	320	0.419	0.264	0.204	0.217	0.255	0.202
0.2	640	0.391	0.252	0.204	0.220	0.261	0.221
0.2	1280	0.382	0.251	0.208	0.227	0.271	0.236
0.4	80	0.764	0.649	0.503	0.387	0.342	0.323
0.4	160	0.793	0.673	0.490	0.382	0.349	0.285
0.4	320	0.864	0.696	0.493	0.391	0.363	0.291
0.4	640	0.905	0.699	0.493	0.393	0.370	0.309
0.4	1280	0.919	0.700	0.492	0.393	0.370	0.319
0.6	80	0.886	0.877	0.805	0.654	0.537	0.470
0.6	160	0.960	0.949	0.837	0.658	0.540	0.450
0.6	320	0.994	0.972	0.841	0.655	0.543	0.455
0.6	640	0.999	0.975	0.838	0.654	0.547	0.462
0.6	1280	1.000	0.976	0.834	0.652	0.548	0.470
0.8	80	0.980	0.989	0.983	0.926	0.811	0.709
0.8	160	0.999	0.999	0.994	0.930	0.809	0.714
0.8	320	1.000	1.000	0.994	0.924	0.802	0.714
0.8	640	1.000	1.000	0.994	0.924	0.800	0.714
0.8	1280	1.000	1.000	0.993	0.920	0.796	0.711

(a) null hypothesis: $\hat{l}_k = 0$;
uses 95% (1.96 standard errors) confidence intervals;
Gaussian random walks, 100,000 iterations each experiment;
model: $\Delta u_t = a\,\Delta p_t + (1-a)e_t$, $\Delta p_t \sim \mathrm{NID}(0,1)$, $e_t \sim \mathrm{NID}(0,1)$;
value of a determines the value of l;
statistic: $1 - p\left[\hat{l}_{kn} - 1.96\sqrt{\frac{2}{3}\frac{k}{T}}\left(1-\hat{l}_{kn}^2\right) < 0 < \hat{l}_{kn} + 1.96\sqrt{\frac{2}{3}\frac{k}{T}}\left(1-\hat{l}_{kn}^2\right)\right]$

Table 4 - Monte Carlo - Zero Long-Run Correlation Test [(a)]

l	T	k/T					
		0.025	0.050	0.100	0.200	0.400	0.800
				Size			
0.0	80	0.542	0.272	0.088	0.025	0.000	0.000
0.0	160	0.306	0.107	0.044	0.022	0.000	0.000
0.0	320	0.137	0.056	0.041	0.025	0.000	0.000
0.0	640	0.067	0.046	0.043	0.029	0.000	0.000
0.0	1280	0.051	0.047	0.045	0.031	0.000	0.000
				Power			
0.2	80	0.605	0.363	0.158	0.050	0.000	0.000
0.2	160	0.460	0.242	0.110	0.044	0.000	0.000
0.2	320	0.382	0.197	0.105	0.050	0.000	0.000
0.2	640	0.350	0.186	0.106	0.052	0.000	0.000
0.2	1280	0.340	0.185	0.109	0.055	0.000	0.000
0.4	80	0.748	0.597	0.365	0.138	0.000	0.000
0.4	160	0.774	0.603	0.334	0.133	0.000	0.000
0.4	320	0.844	0.619	0.336	0.140	0.000	0.000
0.4	640	0.885	0.617	0.333	0.142	0.000	0.000
0.4	1280	0.900	0.618	0.335	0.145	0.000	0.000
0.6	80	0.878	0.852	0.693	0.345	0.000	0.000
0.6	160	0.955	0.927	0.717	0.346	0.000	0.000
0.6	320	0.993	0.955	0.721	0.350	0.000	0.000
0.6	640	0.999	0.960	0.718	0.355	0.000	0.000
0.6	1280	1.000	0.960	0.713	0.355	0.000	0.000
0.8	80	0.978	0.985	0.963	0.742	0.000	0.000
0.8	160	0.999	0.999	0.981	0.747	0.000	0.000
0.8	320	1.000	1.000	0.983	0.742	0.000	0.000
0.8	640	1.000	1.000	0.981	0.740	0.000	0.000
0.8	1280	1.000	1.000	0.980	0.738	0.000	0.000

(a) null hypothesis: $\hat{l}_k = 0$;
theoretical size is 0.05;
Gaussian random walks, 100,000 iterations each experiment;
model: $\Delta u_t = a\Delta p_t + (1-a)e_t$, $\Delta p_t \sim \mathrm{NID}(0,1)$, $e_t \sim \mathrm{NID}(0,1)$;
value of a determines the value of l;
statistic: $1 - p\left[-1.96\sqrt{\frac{2}{3}\frac{k}{T}} < \hat{l}_{kn} < +1.96\sqrt{\frac{2}{3}\frac{k}{T}}\right]$

References

Anderson, T. W. (1971): *The Statistical Analysis of Time Series*. New York: John Wiley & Sons.

Brockwell, P. J. and R. A. Davis (1991): *Time Series: Theory and Methods*. New York: Springer.

Cochrane, J. H. (1988): "How Big is the Random Walk in GNP?," *Journal of Political Economy*, 96, 893-920.

Cochrane, J. H. and A. M. Sbordone (1988): "Multivariate Estimates of the Permanent Components of GNP and Stock Prices," *Journal of Economic Dynamics and Control*, 12, 255-296.

Engle, R. F. and C. W. J. Granger (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251-276.

Fisher, M. E. and J. J. Seater (1993): "Long-Run Neutrality and Superneutrality in an ARIMA Framework," *American Economic Review*, 83, 402-415.

Fuller, W. A. (1976): *Introduction to Statistical Time Series*. New York: John Wiley & Sons.

Granger, C. W. J. and R. Engle (1983): "Applications of Spectral Analysis in Econometrics," in D. R. Brillinger and P. R. Krishnaiah, eds., *Handbook of Statistics*, Vol. 3. Amsterdam: Elsevier.

Granger, C. W. J. and H. J. B. Rees (1968): "Spectral Analysis of the Term Structure of the Interest Rates," *Review of Economic Studies*, 35, 67-76.

Granger, C. W. J. and M. W. Watson (1984): "Time Series and Spectral Methods in Econometrics," in Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics*, Vol. II. New York: North-Holland.

Granger, C. W. J. and A. A. Weiss (1983): "Time Series Analysis of Error-Correction Models," in *Studies in Econometrics, Time Series, and Multivariate Statistics*. New York: Academic Press.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton: Princeton University Press.

Hannan, E. J. (1970): *Multiple Time Series*. New York: John Wiley and Sons.

King, R. G. and M. W. Watson (1994): "The Post-War U.S. Phillips Curve: A Revisionist Econometric History," in A. H. Meltzer and C. I. Plosser, eds., *Carnegie-Rochester Conference Series on Public Policy*, 41. Amsterdam: North-Holland.

King, R. G. and M. W. Watson (1997): "Testing Long-Run Neutrality," *Federal Reserve Bank of Richmond Economic Quarterly*, 83, 3, 69-101.

Koopmans, L. H. (1974): *The Spectral Analysis of Time Series*. San Diego: Academic Press.

McCallum, B. T. (1984): "On Low-Frequency Estimates of Long-Run Relationships in Macroeconomics," *Journal of Monetary Economics*, 14, 3-14.

Nerlove, M. and F. X. Diebold (1990): "Time Series Analysis," in J. Eatwell, M. Milgate and P. Newman, eds., *The New Palgrave: Time Series and Statistics*. New York: W. W. Norton & Company.

Newey, W. K. and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

Newey, W. K. and K. D. West (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

Priestley, M. B. (1981): *Spectral Analysis and Time Series*. London, Academic Press.

Reinsel, G. C. (1993): *Elements of Multivariate Time Series Analysis*. New York: Springer-Verlag.

Sargent, T. J. (1987): *Macroeconomic Theory*. San Diego: Academic Press.

Weber, A. A. (1994): "Testing Long-Run Neutrality: Empirical Evidence for G7-Countries with Special Emphasis on Germany," in A. H. Meltzer and C. I. Plosser, eds., *Carnegie-Rochester Conference Series on Public Policy*, 41, 67-117. Amsterdam: North-Holland.