The New Palgrave Dictionary of Economics Online

serial correlation and serial dependence

Yongmiao Hong
From The New Palgrave Dictionary of Economics, Second Edition, 2008
Edited by Steven N. Durlauf and Lawrence E. Blume

Abstract

In this article we discuss serial correlation in a linear time series regression context and serial dependence in a nonlinear time series context. We first discuss various tests for serial correlation for both estimated regression residuals and observed raw data. Particular attention is paid to the impact of parameter estimation uncertainty and conditional heteroskedasticity on the asymptotic distribution of test statistics. We discuss the drawback of serial correlation in nonlinear time series models and introduce a number of measures that can capture nonlinear serial dependence and reveal useful information about serial dependence.

Keywords

ARMA models; Durbin–Watson statistic; efficient market hypothesis; entropy; generalized spectral density; homoskedasticity; heteroskedasticity; kernel estimators; Lagrange multipliers; rational expectations; serial correlation; serial dependence; spectral density; statistical inference; time series analysis

Article

1 Introduction

Serial correlation and serial dependence have been central to time series econometrics. The existence of serial correlation complicates statistical inference of econometric models; and in time series analysis, inference of serial correlation, or more generally, serial dependence, is crucial to characterize the dynamics of time series processes. Lack of serial correlation is also an important implication of many economic theories and economic hypotheses. For example, the efficient market hypothesis implies that asset returns are an martingale difference sequence (*m.d.s.*), and so are serially uncorrelated. More generally, rational expectations theory implies that the expectational errors of the economic agent are serially uncorrelated. In this article we first discuss various tests for serial correlation, for both estimated model residuals and observed raw data, and we discuss their relationships. We then discuss serial dependence in a nonlinear time series context, introducing related measures and tests for serial dependence.

2 Testing for serial correlation

Consider a linear regression model

$$Y_t = X_t'\beta^0 + \varepsilon_t,\ t=1,\cdots,n,$$

(2.1)

where Y_t is a dependent variable, X_t is a $k\times 1$ vector of explanatory variables, β^0 is an unknown $k\times 1$ parameter vector, and ε_t is an unobservable disturbance with $E(\varepsilon_t|X_t)=0$. Suppose X_t is strictly exogenous such that $\mathrm{cov}(X_t, \varepsilon_s)=0$ for all t, s. Then (2.1) is called a static regression model. If X_t contains lagged dependent variables, (2.1) is called a dynamic regression model.
For a linear dynamic regression model, serial correlation in $\{\varepsilon_t\}$ will generally render inconsistent the OLS estimator. To see this, we consider an AR (1) model

$$Y_t = \beta_0^0 + \beta_1^0 Y_{t-1} + \varepsilon_t = X_t'\beta^0 + \varepsilon_t,$$

where $X_t=(1, Y_{t-1})'$. If $\{\varepsilon_t\}$ also follows an AR(1) process, we will have $E(X_t\varepsilon_t)\neq 0$, rendering inconsistent the OLS estimator for β^0. It is therefore important to check serial correlation for estimated model residuals, which serves as a misspecification test for a linear dynamic regression model. For a static linear regression model, it is also useful to check serial correlation. In particular, if there exists no serial correlation in $\{\varepsilon_t\}$ in a static regression model, then there is no need to use a long-run variance estimator of the OLS estimator $\hat{\beta}$ (for example, Andrews, 1991; Newey and West, 1987).

2.1 Durbin–Watson test

Testing for serial correlation has been a longstanding problem in time series econometrics. The most well known test for serial correlation in regression disturbances is the Durbin–Watson test, which is the first formal procedure developed for testing first order serial correlation

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t,\ \{u_t\} \sim i.i.d.\,(0,\sigma^2)$$

using the OLS residuals $\{e_t\}_{t=1}^n$ in a static linear regression model. Durbin and Watson (1950; 1951) propose a test statistic

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}.$$

Durbin and Watson present tables of bounds at the 0.05, 0.025 and 0.01 significance levels of the d statistic for static regressions with an intercept. Against the one-sided alternative that $\rho>0$, if d is less than the lower bound d_L, the null hypothesis that $\rho=0$ is rejected; if ρ is greater than the upper bound d_U, the null hypothesis is accepted. Otherwise, the test is equivocal. Against the one-sided alternative that $\rho<0$, $4-d$ can be used to replace d in the above procedure.

The Durbin–Watson test has been extended to test for lag 4 autocorrelation by Wallis (1972) and for autocorrelation at any lag by Vinod (1973).

2.2 Durbin's *h* test

The Durbin–Watson d test is not applicable to dynamic linear regression models, because parameter estimation uncertainty in the OLS estimator $\hat{\beta}$ will have nontrivial impact on the distribution of d. Durbin (1970) developed the so-called h test for first-order autocorrelation in $\{\varepsilon_t\}$ that takes into account parameter estimation uncertainty in $\hat{\beta}$. Consider a simple dynamic linear regression model

$$Y_t=\beta_0^0+\beta_1^0 Y_{t-1}+\beta_2^0 X_t+\varepsilon_t,$$

where X_t is strictly exogenous. Durbin's h statistic is defined as:

$$h=\hat{\rho}\sqrt{\frac{n}{1-n\hat{\mathrm{v}}\mathrm{ar}(\hat{\beta}_1)}},$$

where $\hat{\mathrm{v}}\mathrm{ar}(\hat{\beta}_1)$ is an estimator for the asymptotic variance of $\hat{\beta}_1$, $\hat{\rho}$ is the OLS estimator from regressing e_t on e_{t-1} (in fact, $\hat{\rho}\approx 1-d/2$). Durbin (1970) shows that $h\xrightarrow{d}N(0,1)$ as $n\to\infty$ under null hypothesis that $\rho=0$.

2.3 Breusch−Godfrey test

A more convenient and generally applicable test for serial correlation is the Lagrange multiplier test developed by Breusch (1978) and Godfrey (1978). Consider an auxiliary autoregression of order p:

$$\varepsilon_t=\sum_{j=1}^{p}\alpha_j\varepsilon_{t-j}+z_t,\ t=p+1,\cdots,n.$$

(2.2)

The null hypothesis of no serial correlation implies $\alpha_j=0$ for all $1\leq j\leq p$. Under the null hypothesis, we have $nR_{uc}^2\xrightarrow{d}\chi_p^2$, where R_{uc}^2 is the uncentred R^2 of (2.2). However, the autoregression (2.2) is infeasible because ε_t is unobservable. One can replace ε_t with the OLS residual e_t:

$$e_t=\sum_{j=1}^{p}\alpha_j e_{t-j}+v_t,\ t=p+1,\cdots,n.$$

Such a replacement, however, may contaminate the asymptotic distribution of the test statistic because $e_t=\varepsilon_t-(\hat{\beta}-\beta)'X_t$ contains the estimation error $(\hat{\beta}-\beta)'X_t$ where X_t may have nonzero correlation with the regressors e_{t-j} for $1\leq j\leq p$ in dynamic regression models. This correlation affects the asymptotic distribution of nR_{uc}^2 so that it will not be χ_p^2. To purge this impact of the asymptotic distribution of the test statistic, one can consider the augmented auxiliary regression

$$e_t=X_t'\gamma+\sum_{j=1}^{p}\alpha_j e_{t-j}+v_t,\ t=p+1,\cdots,n.$$

(2.3)

The inclusion of X_t will capture the impact of estimation error $(\hat{\beta}-\beta)'X_t$. As a result, the test statistic $nR^2\xrightarrow{d}\chi_p^2$ under the null hypothesis, where, assuming that X_t contains an intercept, R^2 is the centred squared multi-correlation coefficient in (2.3). For a static linear regression model, it is not necessary to include X_t in the auxiliary regression, because $\{X_t\}$ and $\{\varepsilon_t\}$ are uncorrelated, but it does not harm the size of the test if X_t is included. Therefore, the nR^2 test is applicable to both static and dynamic regression models. We note that Durbin's h test is asymptotically equivalent to the nR^2 test of (2.3) with $p=1$.

2.4 Box–Pierce–Ljung test

In time series ARMA modelling, Box and Pierce (1970) propose a portmanteau test as a diagnostic check for the adequacy of an ARMA model

$$Y_t=\psi_0+\sum_{j=1}^{r}\psi_j Y_{t-j}+\sum_{j=1}^{q}\theta_j\varepsilon_{t-j}+\varepsilon_t,\ \left\{\varepsilon_t\right\}\sim i.i.d.(0,\sigma^2).$$

(2.4)

Suppose e_t is an estimated residual obtained from a maximum likelihood estimator. One can define the residual sample autocorrelation function

$$\hat{\rho}(j)=\frac{\hat{\gamma}(j)}{\hat{\gamma}(0)},\ j=0,\ \pm1,\cdots,\ \pm(n-1),$$

where $\hat{\gamma}(j)=n^{-1}\sum_{t=|j|+1}^{n}e_t e_{t-|j|}$ is the residual sample autocovariance function.

Box and Pierce (1970) propose a portmanteau test

$$Q=n\sum_{j=1}^{p}\hat{\rho}^2(j)\xrightarrow{d}\chi_{p-(r+q)}^2,$$

where the asymptotic χ^2 distribution follows under the null hypothesis of no serial correlation, and the adjustment of degrees of freedom $r+q$ is due to the impact of parameter estimation uncertainty for the r autoregressive coefficients and q moving average coefficients in (2.4).

To improve small sample performance of the Q test, Ljung and Box (1978) propose a modified Q test statistic:

$$Q^* \equiv n(n+2)\sum_{j=1}^{p}(n-j)^{-1}\hat{\rho}^2(j) \xrightarrow{d} \chi^2_{p-(r+q)}.$$

The modification matches the first two moments of Q^* with those of the χ^2 distribution. This improves the size in small samples, although not the power of the test.
The Q test is applicable to test serial correlation in the OLS residuals $\{e_t\}$ of a linear static regression model, with $Q \xrightarrow{d} \chi^2_p$ under the null hypothesis. Unlike for ARMA models, there is no need to adjust the degrees of freedom for the χ^2 distribution because the estimation error $(\hat{\beta}-\beta)'X_t$ has no impact on it, due to the fact that cov(X_t, ε_s)=0 for all t, s. In fact, it could be shown that the nR^2 and Q statistics are asymptotically equivalent under the null hypothesis. However, when applied to the estimated residual of a dynamic regression model which contains both endogenous and exogenous variables, the asymptotic distribution of the Q test is generally unknown (Breusch and Pagan, 1980). One solution is to modify the Q test statistic as follows:

$$\hat{Q} \equiv n\hat{\rho}'(I-\hat{\Phi})^{-1}\hat{\rho} \xrightarrow{d} \chi^2_p \text{ as } n \to \infty,$$

where $\hat{\rho} = [\hat{\rho}(1), \cdots, \hat{\rho}(p)]'$ and $\hat{\Phi}$ captures the impact caused by nonzero correlation between $\{X_t\}$ and $\{\varepsilon_s\}$. See Hayashi (2000, Section 2.10) for more discussion.

2.5 Spectral density-based test

Much criticism has been levelled at the possible low power of the Box–Pierce–Ljung portmanteau tests, which also applies to the nR^2 test, due to the asymptotic equivalence between the Q test and the nR^2 test for a static regression. Moreover, there is no theoretical guidance on the choice of p for these tests. A fixed lag order p will render inconsistent any test for serial correlation of unknown form.
To test serial correlation of unknown form in the estimated residuals of a linear regression model, which can be static or dynamic, Hong (1996) uses a kernel spectral density estimator

$$\hat{h}(\omega) = \frac{1}{2\pi}\sum_{j=1-n}^{n-1} k(j/p)\hat{\gamma}(j)e^{-ij\omega},\ \omega \in [-\pi, \pi],$$

and compares it with the flat spectrum implied by the null hypothesis of no serial correlation:

$$\hat{h}_0(\omega) = \frac{1}{2\pi}\hat{\gamma}(0),\ \omega \in [-\pi, \pi].$$

Under the null hypothesis, $\hat{h}(\omega)$ and $\hat{h}_0(\omega)$ are close. If $\hat{h}(\omega)$ is significantly different from $\hat{h}_0(\omega)$ there is evidence of serial correlation. A global measure of the divergence between $\hat{h}(\omega)$ and $\hat{h}_0(\omega)$ is the quadratic form

$$L^2(\hat{h}, \hat{h}_0) \equiv \int_{-\pi}^{\pi}\left[\hat{h}(\omega) - \hat{h}_0(\omega)\right]^2 d\omega = \sum_{j=1}^{n-1} k^2(j/p)\hat{\gamma}^2(j).$$

The test statistic is a normalized version of the quadratic form:

$$M_o \equiv \left[n\sum_{j=1}^{n-1} k^2(j/p)\hat{\rho}^2(j) - \hat{C}_o(p)\right] / \sqrt{\hat{D}_o(p)} \xrightarrow{d} N(0,1)$$

where the centring and scaling factors

$$\hat{C}_o(p) = \sum_{j=1}^{n-1}(1-j/n)k^2(j/p),$$

$$\hat{D}_o(p) = 2\sum_{j=1}^{n-2}(1-j/n)[1-(j+1)/n]k^4(j/p).$$

This test can be viewed as a generalized version of Box and Pierce's (1970) portmanteau test, the latter being equivalent to using the truncated kernel $k(z)=\mathbf{1}(|z|\leq 1)$, which gives equal weighting to each of the first p lags. In this case, M_o is asymptotically equivalent to

$$M_T \equiv \frac{n\sum_{j=1}^{p}\hat{\rho}^2(j) - p}{\sqrt{2p}} \xrightarrow{d} \frac{\chi^2_p - p}{\sqrt{2p}} \sim N(0,1) \text{ as } p \to \infty.$$

However, uniform weighting to different lags may not be powerful when a large number of lags is employed. For any weakly stationary process, the autocovariance function $\gamma(j)$ typically decays to 0 as lag order j increases. Thus, it is more efficient to discount higher order lags. This can be achieved by using non-uniform kernels. Most commonly used kernels, such as the Bartlett, Pazren and quadratic-spectral kernels, discount higher order lags. Hong (1996) shows that the Daniell kernel $k(z)=\sin(\pi z)/(\pi z)$, $-\infty<z<\infty$, maximizes the power of the M test over a wide class of the kernel functions when $p\to\infty$. The optimal kernel for hypothesis testing differs from the optimal kernel for spectral density estimation.
It is important to note that the spectral density test M applies to both static and dynamic regression models, and no modification is needed when applied to a dynamic regression model. Intuitively, parameter estimation uncertainty causes some adjustment of degrees of freedom, which becomes asymptotically independent when the lag order $p\to\infty$ as $n\to\infty$. This differs from the case where p is fixed.
For similar spectral density-based tests for serial correlation, see Paparoditis (2000), Chen and Deo (2004), and Fan and Zhang (2004).

2.6 Heteroskedasticity–robust tests

All the aforementioned tests assume conditional homoskedasticity or even *i.i.d.* on $\{\varepsilon_t\}$. This rules out high frequency financial time series, which have been documented to have persistent volatility clustering. Some effort has been devoted to robustifying tests for serial correlation. Wooldridge

(1990; 1991) proposes a two-stage procedure to robustify the nR^2 test for serial correlation in estimated residuals $\{e_t\}$ of a linear regression model (2.1): (i) regress $(e_{t-1},\ldots,e_{t-p})$ on X_t and save the estimated $p\times 1$ residual vector $\hat{v}_t$; (ii) regress 1 on $\hat{v}_t e_t$ and obtain SSR, the sum of squared residuals; (iii) compare the $n-SSR$ statistic with the asymptotic χ_p^2 distribution. The first auxiliary regression purges the impact of parameter estimation uncertainty in the OLS estimator $\hat{\beta}$ and the second auxiliary regression delivers a test statistic robust to conditional heteroskedasticity of unknown form.

Whang (1998) also proposes a semiparametric test for serial correlation in estimated residuals of a possibly nonlinear regression model. Assuming that $\varepsilon_t=\sigma[Z_t(\alpha)]z_t$, where $\{z_t\}$~i.i.d.(0,1), $\mathrm{var}(\varepsilon_t|I_{t-1})=\sigma^2[Z_t(\alpha)]$ depends on a random vector with fixed dimension (for example, $Z_t(\alpha)=(\varepsilon_{t-1}^2,\ldots,\varepsilon_{t-K}^2)'$ for a fixed K), but the functional form $\sigma^2(\cdot)$ is unknown. This covers a variety of conditionally heteroskedastic processes, although it rules out non-Markovian processes such as Bollerslev's (1986) GARCH model. Whang (1998) first estimates $\sigma^2[Z_t(\alpha)]$ using a kernel method, and then constructs a Box–Pierce type test for serial correlation in the estimated regression residuals standardized by the square root of the nonparametric variance estimator.

The assumption imposed on $\mathrm{var}(\varepsilon_t|I_{t-1})$ in Whang (1998) rules out GARCH models, and both Wooldridge (1991) and Whang (1998) test serial correlation up to a fixed lag order only. Hong and Lee (2007) have recently robustified Hong's (1996) spectral density-based consistent test for serial correlation of unknown form:

$$\hat{M}\equiv\left[n^{-1}\sum_{j=1}^{n-1}k^2(j/p)\hat{\gamma}^2(j)-\hat{C}(p)\right]/\sqrt{\hat{D}(p)},$$

where the centring and scaling factors

$$\hat{C}(p)\equiv\hat{\gamma}^2(0)\sum_{j=1}^{n-1}(1-j/n)k^2(j/p)+\sum_{j=1}^{n-1}k^2(j/p)\hat{\gamma}_{22}(j),$$

$$\hat{D}(p)\equiv 2\hat{\gamma}^4(0)\sum_{j=1}^{n-2}(1-j/n)[1-(j+1)/n]k^4(j/p)+4\hat{\gamma}^2(0)\sum_{j=1}^{n-2}k^4(j/p)\hat{\gamma}_{22}(j)+2\sum_{j=1}^{n-2}\sum_{l=1}^{n-2}k^2(j/p)k^2(l/p)\hat{C}^2(0,j,l),$$

with $\hat{\gamma}_{22}(j)\equiv n^{-1}\sum_{t=j+1}^{n-1}[e_t^2-\hat{\gamma}(0)][e_{t-j}^2-\hat{\gamma}(0)]$ and $\hat{C}(0,j,l)\equiv n^{-1}\sum_{t=\max(j,l)+1}^{n}[e_t^2-\hat{\gamma}(0)]e_{t-j}e_{t-l}$. Intuitively, the centring and scaling factors have taken into account possible volatility clustering and asymmetric features of volatility dynamics, so the $\hat{M}$ test is robust to these effects. It allows for various volatility processes, including GARCH models, Nelson's (1991) EGARCH, and Glosten, Jagannathan and Runkle's (1993) Threshold GARCH models.

Martingale tests

Several tests for serial correlation are motivated for testing the *m.d.s.* property of an observed time series $\{Y_t\}$, say asset returns, rather than estimated residuals of a regression model. We now present a unified framework to view some martingale tests for observed data.

Extending an idea of Cochrane (1988), Lo and MacKinlay (1988) first rigorously present an asymptotic theory for a variance ratio test for the *m.d.s.* hypothesis of $\{Y_t\}$. Because the *m.d.s.* hypothesis implies $\gamma(j)=0$ for all $j>0$, one has

$$\frac{\mathrm{var}\left(\sum_{j=1}^{p}Y_{t-j}\right)}{p\cdot\mathrm{var}(Y_t)}=\frac{p\gamma(0)+2p\sum_{j=1}^{p}(1-j/p)\gamma(j)}{p\gamma(0)}=1.$$

This unity property of the variance ratio can be used to test the *m.d.s.* hypothesis because any departure from unity is evidence against the *m.d.s.* hypothesis.

The variance ratio test is essentially based on the statistic

$$\mathrm{VR}_0\equiv\sqrt{n/p}\sum_{j=1}^{p}(1-j/p)\hat{\rho}(j)=\frac{\pi}{2}\sqrt{n/p}\left[\hat{f}(0)-\frac{1}{2\pi}\right],$$

where $\hat{f}(0)$ is a kernel-based normalized spectral density estimator at frequency 0, with the Bartlett kernel $k(z)=(1-|z|)\,\mathbf{1}(|z|\le 1)$ and a lag order p. In other words, VR_o is based on a spectral density estimator of frequency 0, and because of this, it is particularly powerful against long memory processes, whose spectral density at frequency 0 is infinity (see Robinson, 1994, for an excellent survey).

Under the *m.d.s.* hypothesis with conditional homoskedasticity, Lo and MacKinlay (1988) show that for any fixed p,

$$\mathrm{VR}_0\xrightarrow{d}N[0,2(2p-1)(p-1)/3p]\ \text{as}\ n\to\infty.$$

Lo and MacKinlay (1988) also consider a heteroskedasticity-consistent variance ratio test:

$$\mathrm{VR}\equiv\sqrt{n/p}\sum_{j=1}^{p}(1-j/p)\hat{\gamma}(j)/\sqrt{\hat{\gamma}_2(j)},$$

where $\hat{\gamma}_2(j)$ is a consistent estimator for the asymptotic variance of $\hat{\gamma}(j)$ under conditional heteroskedasticity. Lo and MacKinlay (1988) assume a fourth order cumulant condition that

$$E[(Y_t-\mu)^2(Y_{t-l}-\mu)(Y_{t-1}-\mu)]=0,\ j,l>0,\ j\neq l$$
(2.5)

Intuitively, this condition ensures that the sample autocovariances at different lags are asymptotically uncorrelated; that is, $\mathrm{cov}\left[\sqrt{n}\hat{\gamma}(j),\sqrt{n}\hat{\gamma}(l)\right]\to 0$ for all $j\neq l$. As a result, the heroskedasticity-consistent VR has the same asymptotic distribution as VR_o. However, the condition in (2.5) rules out many important volatility processes, such as EGARCH and Threshold GARCH models. Moreover, the variance ratio test only exploits the implication of the *m.d.s.* hypothesis on the spectral density at frequency 0; it does not check the spectral density at nonzero frequencies. As a result, it is not consistent against serial correlation of unknown form. See Durlauf (1991) for more discussion.

Durlauf (1991) considers testing the *m.d.s.* hypothesis for observed raw data $\{Y_t\}$, using the spectral distribution function

$$H(\lambda) \equiv 2\int_0^{\pi\lambda} h(\omega)\,d\omega = \gamma(0)\lambda + \sqrt{2}\sum_{j=1}^{\infty}\gamma(j)\frac{\sqrt{2}\sin(j\pi\lambda)}{j\pi},\ \lambda\in[0,1],$$

where $h(\omega)$ is the spectral density of $\{Y_t\}$:

$$h(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)\cos(j\omega),\ \omega\in[-\pi,\pi].$$

Under the *m.d.s.* hypothesis, $H(\lambda)$ becomes a straight line:

$$H_0(\lambda) = \gamma(0)\lambda,\ \ \lambda\in[0,1].$$

An *m.d.s.* test can be obtained by comparing a consistent estimator for $H(\lambda)$ and $\hat{H}_0(\lambda) = \hat{\gamma}(0)\lambda$.
Although the periodogram (or sample spectral density function)

$$\hat{I}(\omega) \equiv \frac{1}{2\pi n}\left|\sum_{t=1}^{n}(Y_t-\bar{Y})e^{it\omega}\right|^2 = \frac{1}{2\pi}\sum_{j=1-n}^{n-1}\hat{\gamma}(j)e^{-ij\omega}$$

is not consistent for the spectral density $h(\omega)$, the integrated periodogram

$$\hat{H}(\lambda) \equiv 2\int_0^{\lambda\pi}\hat{I}(\omega)\,d\omega = \hat{\gamma}(0)\lambda + \sqrt{2}\sum_{j=1}^{n-1}\hat{\gamma}(j)\frac{\sqrt{2}\sin(j\pi\lambda)}{j\pi}$$

is consistent for $H(\lambda)$, thanks to the smoothing provided by the integration. Among other things, Durlauf (1991) proposes a Cramer–von Mises type statistic

$$CVM \equiv \frac{1}{2}n\int_0^1\left[\hat{H}(\lambda)/\hat{\gamma}(0)-\lambda\right]^2 d\lambda = n\sum_{j=1}^{n-1}\hat{\rho}^2(j)/(j\pi)^2.$$

Under the *m.d.s.* hypothesis with conditional homoskedasticity, Durlauf (1991) shows

$$CVM \xrightarrow{d} \sum_{j=1}^{\infty}\chi_j^2(1)/(j\pi)^2,$$

where $\left\{\chi_j^2(1)\right\}_{j=1}^{\infty}$ is a sequence of *i. i. d.* χ^2 random variables with one degree of freedom. This asymptotic distribution is nonstandard, but it is distribution-free and can be easily tabulated or simulated. An appealing property of Durlauf's (1991) test is its consistency against serial correlation of unknown form, and there is no need to choose a lag order p.
Deo (2000) shows that under the *m.d.s.* hypothesis with conditional heteroskedasticity, Durlauf's (1991) test statistic can be robustified as follows:

$$CVM = \sum_{j=1}^{n-1}\frac{\left[\hat{\gamma}^2(j)/\hat{\gamma}_2(j)\right]}{(j\pi)^2} \xrightarrow{d} \sum_{j=1}^{\infty}\chi_j^2(1)/(j\pi)^2.$$

where $\hat{\gamma}_2(j)$ is a consistent estimator for the asymptotic variance of $\hat{\gamma}(j)$ and the asymptotic distribution remains unchanged. Like Lo and MacKinlay (1988), Deo (2000) also imposes the crucial fourth order joint cumulant condition in (2.5).

3 Serial dependence in nonlinear models

The autocorrelation function $\gamma(j)$, or equivalently, the power spectrum $h(\omega)$, of a time series $\{Y_t\}$, is a measure for linear association. When $\{Y_t\}$ is a stationary Gaussian process, $\gamma(j)$ or $h(\omega)$ can completely determine the full dynamics of $\{Y_t\}$.
It has been well documented, however, that most economic and financial time series, particularly high-frequency economic and financial time series, are not Gaussian. For non-Gaussian processes, $\gamma(j)$ and $h(\omega)$ may not capture the full dynamics of $\{Y_t\}$. We consider two nonlinear process examples:

- Bilinear (BL) autoregressive process:
  $$Y_t = \alpha\varepsilon_{t-1}Y_{t-2} + \varepsilon_t,\ \left\{\varepsilon_t\right\} \sim i.\,i.\,d.\,(0,\sigma^2).$$
  (3.1)
- Nonlinear moving average (NMA) process:
  $$Y_t = \alpha\varepsilon_{t-1}\varepsilon_{t-2} + \varepsilon_t,\ \left\{\varepsilon_t\right\} \sim i.\,i.\,d.\,(0,\sigma^2).$$
  (3.2)

For these two processes, there exists nonlinearity in conditional mean: $E(Y_t|I_{t-1})=\alpha\varepsilon_{t-1}Y_{t-2}$ under (3.1) and $E(Y_t|I_{t-1})=\alpha\varepsilon_{t-1}Y_{t-2}$ under (3.2). However, both processes are serially uncorrelated. If $\{Y_t\}$ follows either a BL process in (3.1) or a NMA process in (3.2), $\{Y_t\}$ is not *m.d.s.* but $\gamma(j)$ and $h(\omega)$ will miss it. Hong and Lee (2003a) document that indeed, for foreign currency markets, most foreign exchange changes are serially uncorrelated, but they are all not *m.d.s.* There exist predictable nonlinear components in the conditional mean of foreign exchange markets.
Serial dependence may also exist only in higher order conditional moments. An example is Engle's (1982) first order autoregressive conditional heteroskedastic (ARCH (1)) process:

$$\begin{cases} Y_t = \sigma_t\varepsilon_t, \\ \sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2, \\ \{\varepsilon_t\} \sim i.\,i.\,d.\,(0,1). \end{cases}$$
(3.3)

For this process, the conditional mean $E(Y_t|I_{t-1})=0$; which implies $\gamma(j)=0$ for all $j>0$.

However, the conditional variance, $var(Y_t|I_{t-1}) = \alpha_0 + \alpha_1 Y_{t-1}^2$, depends on the previous volatility. Both $\gamma(j)$ and $h(\omega)$ will miss such higher order dependence.

In nonlinear time series modelling, it is important to measure serial dependence, that is, any departure from *i.i.d.*, rather than merely serial correlation. As Priestley (1988) points out, the main purpose of nonlinear time series analysis is to find a filter $h(\cdot)$ such that

$$h(Y_t, Y_{t-1}, \cdots) = \varepsilon_t \sim i.i.d.(0, \sigma^2).$$

In other words, the filter $h(\cdot)$ can capture all serial dependence in $\{Y_t\}$ so that the 'residual' $\{\varepsilon_t\}$ becomes an *i.i.d.* sequence. One example of $h(\cdot)$ in modelling the conditional probability distribution of Y_t given I_{t-1}, is the probability integral transform

$$Z_t(\beta) = \int_{-\infty}^{Y_t} f(y|I_{t-1}, \beta)\, dy,$$

where $f(y|I_{t-1}, \beta)$ is a conditional density model for Y_t given and I_{t-1}, and β is an unknown parameter. When $f(y|I_{t-1}, \beta)$ is correctly specified for the conditional probability density of Y_t given I_{t-1}, that is, when the true conditional density coincides with $f(y|I_{t-1}, \beta^0)$ for some β^0, the probability integral transforms becomes

$$\left\{Z_t(\beta^0)\right\} \sim i.i.d.\,\mathrm{U}[0,1].$$

(3.4)

Thus, one can test whether $f(y|I_{t-1}, \beta)$ is correctly specified by checking the *i.i.d.*U[0,1] for the probability integral transform series.

3.1 Bispectrum and higher-order spectra

Because the autocorrelation function $\gamma(j)$ and the spectral density $h(\omega)$ are rather limited in nonlinear time series analysis, various alternative tools have been proposed to capture nonlinear serial dependence (for example, Granger and Terasvirta, 1993; Tjøstheim, 1996). For example, one often uses the third-order cumulant function

$$C(j, k) \equiv E[(Y_t - \mu)(Y_{t-j} - \mu)(Y_{t-k} - \mu)],\; j, k = 0, \pm 1, \cdots.$$

This is also called the biautocovariance function of $\{Y_t\}$. It can capture certain nonlinear time series, particularly those displaying asymmetric behaviours such as skewness. Hsieh (1989) proposes a test based on $C(j, k)$ for a given pair of (j, k) which can detect some predictable nonlinear components in asset returns.

The Fourier transform of $C(j, k)$,

$$b(\omega_1, \omega_2) \equiv \frac{1}{(2\pi)^2} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} C(j, k) e^{-ij\omega_1 - ik\omega_2},\; \omega_1, \omega_2 \in [-\pi, \pi],$$

is called the bispectrum of $\{Y_t\}$. When $\{Y_t\}$ is *i.i.d.*, $b(\omega_1, \omega_2)$ becomes a flat bispectral surface:

$$b_0(\omega_1, \omega_2) \equiv \frac{E(Y_t^3)}{(2\pi)^2},\; \omega_1, \omega_2 \in [-\pi, \pi].$$

Any deviation from a flat bispectral surface will indicate the existence of serial dependence in $\{Y_t\}$. Moreover, $b(\omega_1, \omega_2)$ can be used to distinguish some linear time series processes from nonlinear time series processes. When $\{Y_t\}$ is a linear process with *i.i.d.* innovations, that is, when

$$Y_t = \alpha_0 + \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j} + \varepsilon_t,\; \left\{\varepsilon_t\right\} \sim i.i.d.(0, \sigma^2),$$

the normalized bispectrum

$$|\tilde{b}(\omega_1, \omega_2)|^2 \equiv \frac{|b(\omega_1, \omega_2)|^2}{h(\omega_1)h(\omega_2)h(\omega_1 + \omega_2)} = \frac{[E(\varepsilon_t^3)]^2}{2\pi\sigma^6}$$

is a flat surface. Any departure from a flat normalized bispectral surface will indicate that $\{Y_t\}$ is not a linear time series with *i.i.d.* innovations.

The bispectrum $b(\omega_1, \omega_2)$ can capture the BL and NMA processes in (3.1) and (3.2), because the third order cumulant $C(j, k)$ can distinguish them from an *i.i.d* process. However, it may still miss some important alternatives. For example, it will easily miss ARCH (1) with *i.i.d.* N(0,1) innovation $\{\varepsilon_t\}$. In this case, $b(\omega_1, \omega_2)$ becomes a flat bispectrum and cannot distinguish ARCH (1) from an *i.i.d.* sequence. One could use higher order spectra or polyspectra (Brillinger and Rosenblatt, 1967a; 1967b), which are the Fourier transforms of higher order cumulants. However, higher-order spectra have met with some difficulty in practice: Their spectral shapes are difficult to interpret, and their estimation is not stable in finite samples, due to the assumption of the existence of higher order moments. Indeed, it is often a question whether economic and financial data, particularly high-frequency data, have finite higher order moments.

3.2 Nonparametric measures of serial dependence

Nonparametric measures for serial dependence have been proposed in the literature, which avoid assuming the existence of moments. Granger and Lin (1994) propose a nonparametric entropy measure for serial dependence to identify significant lags in nonlinear time series. Define the Kullback–Leibler information criterion

$$I(j) = \int \ln\left[\frac{f_j(x, y)}{g(x)g(y)}\right] f_j(x, y)\, dx dy,\; j = 1, 2, \ldots$$

where $f_j(x, y)$ is the joint probability density of Y_t and Y_{t-j}, and $g(x)$ is the marginal probability density of $\{Y_t\}$. The Granger–Lin normalized entropy measure is defined as follows:

$$e^2(j) = 1 - \exp[-2I(j)],$$

which enjoys some appealing features. For example, $e(j)=0$ if and only if Y_t and Y_{t-j} are independent, and it is invariant to any monotonic continuous transformation. Because $f_j(x, y)$ and $g(x)$ are unknown, Granger and Lin (1994) use nonparametric kernel density estimators. They establish the consistency of their entropy estimator (say $\hat{I}(j)$) but do not derive its asymptotic distribution, which is important for confidence interval estimation and hypothesis testing.

In fact, Robinson (1991) has elegantly explained the difficulty of obtaining the asymptotic distribution of $\hat{I}(j)$ for serial dependence, namely it is a degenerate statistic so that the usual root-n normalization does not deliver a well-defined asymptotic distribution. Robinson (1991) considers a modified entropy estimator

$$\hat{I}_\gamma(j) = n^{-1}\sum_{t=j+1}^{n} C_t(\gamma)\ln\left[\frac{\hat{f}_j(Y_t, Y_{t-j})}{\hat{g}(Y_t)\hat{g}(Y_{t-j})}\right],$$

where $\hat{f}_j(\cdot,\cdot)$ and $\hat{g}(\cdot)$ are nonparametric kernel density estimators, $C_t(\gamma)=1-\gamma$ if t is odd, $C_t(\gamma)=1+\gamma$ if t is even, and γ is a pre-specified parameter. The weighting device $C_t(\gamma)$ does not affect the consistency of $\hat{I}_\gamma(j)$ to $I(j)$ and affords a well-defined asymptotic N(0,1) distribution under the *i.i.d.* hypothesis.

Skaug and Tjøstheim (1993a; 1996) use a different weighting function to avoid the degeneracy of the entropy estimator for serial dependence:

$$\hat{I}_W(j) = n^{-1}\sum_{t=1}^{n} w(Y_t, Y_{t-j})\ln\left[\frac{\hat{f}_j(Y_t, Y_{t-j})}{\hat{g}(Y_t)\hat{g}(Y_{t-j})}\right],$$

where $w(Y_t, Y_{t-j})$ is a weighting function of observations X_t and X_{t-j}. Unlike using Robinson's (1991) weighting device, $\hat{I}_W(j)$ is not consistent for the population entropy $I(j)$, but it also delivers a well-defined asymptotic N(0, 1) distribution after a root-n normalization.

Intuitively, the use of weighting devices slows down the convergence rate of the entropy estimators, giving a well-defined asymptotic N(0,1) distribution after the usual root-n normalization. However, this is achieved at the cost of an efficiency loss, due to the slower convergence rate. Moreover, this approach breaks down when $\{Y_t\}$ is uniformly distributed, as in the case of the probability integral transforms of the conditional density in (3.4). Instead of using a weighting device, Hong and White (2005) exploit the degeneracy of $\hat{I}(j)$ and use a degenerate U-statistic theory to establish its asymptotic normality. Specifically, Hong and White (2005) show

$$nh\hat{I}(j) + hd_n^0 \xrightarrow{d} N(0, V),$$

where $h=h(n)$ is the bandwidth, and d_n^0 and V are nonstochastic factors. This approach preserves the convergence rate of the unweighted entropy estimator, giving sharper confidence interval estimation and more powerful hypothesis tests. It is applicable when $\{Y_t\}$ is uniformly distributed.

Skaug and Tjøstheim (1993b) also use an Hoeffding measure to test serial dependence (see also Delgado, 1996; Hong, 1998; 2000). The empirical Hoeffding measures are based on the empirical distribution functions, which avoid smoothed nonparametric density estimation.

3.3 Generalized spectrum

Without assuming the existence of higher order moments, Hong (1999) proposes a generalized spectrum as an alternative analytic tool to the power spectrum and higher order spectra. The basic idea is to transform $\{Y_t\}$ via a complex-valued exponential function

$$Y_t \to \exp(iuY_t),\ u\in(-\infty, \infty),$$

and then consider the spectrum of the transformed series. Let $\psi(u) = E(e^{iuY_t})$ be the marginal characteristic function of $\{Y_t\}$ and let $\psi_j(u, v) = E[e^{i(uY_t+vY_{t-|j|})}],\ j=0, \pm 1, \dots$ be the pairwise joint characteristic function of. $(Y_t, Y_{t-|j|})$. Define the covariance function between transformed variables e^{iuY_t} and $e^{ivY_{t-|j|}}$:

$$\sigma_j(u, v) = \mathrm{cov}(e^{iuY_t}, e^{ivY_{t-|j|}}),\ j=0, \pm 1, \dots$$

Straightforward algebra yields $\sigma_j(u, v)=\psi_j(u, v)-\psi(u)\,\psi(v)$, which is zero for all u, v if and only if Y_t and $Y_{t-|j|}$ are independent. Thus $\sigma_j(u, v)$ can capture any type of pairwise serial dependence over various lags, including those with zero autocorrelation. For example, $\sigma_j(u, v)$ can capture the BL, NMA and ARCH (1) processes in (3.1)–(3.3), all of which are serially uncorrelated.

The Fourier transform of the generalized covariance $\sigma_j(u, v)$:

$$f(\omega, u, v) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sigma_j(u, v)e^{-ij\omega},\ \omega\in[-\pi, \pi],$$

is called the 'generalized spectral density' of $\{Y_t\}$. Like $\sigma_j(u, v)$, $f(\omega, u, v)$ can capture any type of pairwise serial dependencies in $\{Y_t\}$ over various lags. Unlike the power spectrum and higher order spectra, $f(\omega, u, v)$ does not require any moment condition on $\{Y_t\}$. When var(Y_t) exists, the power spectrum of $\{Y_t\}$ can be obtained by differentiating $f(\omega, u, v)$ with respect to $(u\ v)$ at (0, 0):

$$h(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)e^{-ij\omega} = -\frac{\partial^2}{\partial u\,\partial v} f(\omega, u, v)|_{(u,v)=(0,0)},\ \omega\in[-\pi, \pi].$$

This is the reason why $f(\omega, u, v)$ is called the 'generalized spectral density' of $\{Y_t\}$.

When $\{Y_t\}$ is *i.i.d.*, $f(\omega, u, v)$ becomes a flat generalized spectrum as a function of ω:

$$f_0(\omega, u, v) = \frac{1}{2\pi}\sigma_0(u, v),\ \omega\in[-\pi, \pi].$$

Any deviation of $f(\omega, u, v)$ from the flat generalized spectrum $f_0(\omega, u, v)$ is evidence of serial dependence. Thus, $f(\omega, u, v)$ is suitable to capture any departures from *i.i.d.* Hong and Lee (2003b) use the generalized spectrum to develop a test for the adequacy of nonlinear time series models by checking whether the standardized model residuals are *i.i.d.* Tests for *i.i.d.* are more suitable than tests for serial correlation in nonlinear contexts. Indeed, Hong and Lee (2003b) find that some popular EGARCH models are inadequate in capturing the full dynamics of stock returns, although the

standardized model residuals are serially uncorrelated.
Insight into the ability of $f(\omega, u, v)$ can be gained by considering a Taylor series expansion
$$f(\omega, u, v) = \sum_{m=0}^{\infty}\sum_{l=0}^{\infty} \frac{(iu)^m (iv)^l}{m!l!}\left[\frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \operatorname{cov}(X_t^m, X_{t-|j|}^l)e^{-ij\omega}\right].$$

Although $f(\omega, u, v)$ has no physical interpretation, it can be used to characterize cyclical movements caused by linear and nonlinear serial dependence. Examples of nonlinear cyclical movements include cyclical volatility clustering, and cyclical distributional tail clustering (for example, Engle and Manganelli's (2004) CAVaR model). Intuitively, the supremum function
$$s(\omega) = \sup_{-\infty<u,v<\infty} |f(\omega, u, v)|, \omega \in [-\pi, \pi],$$

can measure the maximum dependence at frequency ω of $\{Y_t\}$. It can be viewed as an operational frequency domain analogue of Granger and Terasvirta's (1993) maximum correlation measure
$$mm_\rho(j) = \max_{g(\cdot), h(\cdot)} |corr[g(Y_t), h(X_{t-j})]|.$$

Once generic serial dependence is detected using $f(\omega, u, v)$ or any other dependence measure, one may like to explore the nature and pattern of serial dependence. For example, one may be interested in the following questions:

- Is serial dependence operative primarily through the conditional mean or through conditional higher order moments?
- If serial dependence exists in conditional mean, is it linear or nonlinear?
- If serial dependence exists in conditional variance, does there exist linear or nonlinear and asymmetric ARCH?

Different types of serial dependence have different economic implications. For example, the efficient market hypothesis fails if and only if there is no serial dependence in conditional mean.
Just as the characteristic function can be differentiated to generate various moments, generalized spectral derivatives, when they exist, can capture various specific aspects of serial dependence, thus providing information on possible types of serial dependence. Suppose $E[(Y_t)^{2\max(m,l)}] < \infty$ for some nonnegative integers m, l. Then the following generalized spectral derivative exists:
$$f^{(0,m,l)}(\omega, u, v) = \frac{\partial^{m+l}}{\partial u^m \partial v^l} f(\omega, u, v) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \sigma_j^{(m,l)}(u, v)e^{-ij\omega},$$

where $\sigma_j^{(m,l)}(u, v) \equiv \partial^{m+l}\sigma_j(u, v) / \partial u^m \partial v^l$. As an illustrative example, we consider the generalized spectral derivative of order (m, l)=(1,0):
$$f^{(0,1,0)}(\omega, u, v) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \sigma_j^{(1,0)}(u, v)e^{-ij\omega}.$$

Observe $\sigma_j^{(1,0)}(0, v) \equiv \operatorname{cov}(iY_t, e^{ivY_{t-|j|}}) = 0$ for all $v \in (-\infty,\infty)$ if and only if $E(Y_t|Y_{t-|j|})=E(Y_t)$ a.s. The function $E(Y_t|Y_{t-|j|})$ is called the autoregression function of $\{Y_t\}$ at lag j. It can capture a variety of linear and nonlinear dependencies in conditional mean, including the BL and NMA processes in (3.1) and (3.2). (The use of $\sigma_j^{(1,0)}(0, v)$, which can be easily estimated by a sample average, avoids smoothed nonparametric estimation of $E(Y_t|Y_{t-|j|})$.) Thus, the generalized spectral derivative $f^{(0,1,0)}(\omega, u, v)$ can be used to capture a wide range of serial dependence in conditional mean. In particular, the function
$$s(\omega) = \sup_{-\infty<v<+\infty} |f^{(0,1,0)}(\omega, 0, v)|$$

can be viewed as an operational frequency domain analogue of Granger and Terasvirta's (1993) maximum mean correlation measure
$$mm(j) = \max_{h(\cdot)}|\operatorname{corr}(Y_t, h(Y_{t-j})|.$$

See Hong and Lee (2005) for more discussion.
Suppose one has found evidence of serial dependence in conditional mean using $f^{(0,1,0)}(\omega, u, v)$ or any other suitable measure, one can go further to explore whether there exists linear serial dependence in mean. This can be done by using the (1,1)-th order generalized derivative
$$f^{(0,1,1)}(\omega, 0, 0) = -h(\omega),$$

which checks serial correlation. Moreover, one can further use $f^{(0,1,l)}(\omega, u, v)$ for $l \geq 2$ to reveal nonlinear serial dependence in mean. In particular, these higher-order derivatives can suggest that there exist: (i) an ARCH-in-mean effect (for example, Engle, Lilien and Robins, 1987) if $\operatorname{cov}(Y_t, Y_{t-j}^2) \neq 0$, (ii) a skewness-in-mean effect (for example, Harvey and Siddique, 2000) if $\operatorname{cov}(Y_t, Y_{t-j}^3) \neq 0$, and (iii) kurtosis-in-mean effect (for example, Brooks, Burke and Persand, 2005) if $\operatorname{cov}(Y_t, Y_{t-j}^4) \neq 0$. These effects may arise from the existence of a time-varying risk premium, asymmetry of market behaviours, and inadequate account for large losses, respectively.

See Also

- kernel estimators in econometrics
- spectral analysis

I thank Steven Durlauf (editor) for suggesting this topic and comments on an earlier version, and Jing Liu for excellent research assistance and references. This research is supported by the Cheung Kong Scholarship of the Chinese Ministry of Education and Xiamen University. All remaining errors are solely mine.

Bibliography

Andrews, D.W.K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–58.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedastcity. *Journal of Econometrics* 31, 307–27.

Box, G.E.P. and Pierce, D.A. 1970. Distribution of residual autocorrelations in autoregressive moving average time series models. *Journal of the American Statistical Association* 65, 1509–26.

Breusch, T.S. 1978. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers* 17, 334–55.

Breusch, T.S. and Pagan, A. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47, 239–53.

Brillinger, D.R. and Rosenblatt, M. 1967a. Asymptotic theory of estimates of kth order spectra. In *Spectral Analysis of Time Series*, ed. B. Harris. New York: Wiley.

Brillinger, D.R. and Rosenblatt, M. 1967b. Computation and interpretation of the kth order spectra. In *Spectral Analysis of Time Series*, ed. B. Harris. New York: Wiley.

Brooks, C., Burke, S. and Persand, G. 2005. Autoregressive conditional kurtosis. *Journal of Financial Econometrics* 3, 399–421.

Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Chen, W. and Deo, R. 2004. A generalized portmanteau goodness-of-fit test for time series models. *Econometric Theory* 20, 382–416.

Cochrane, J.H. 1988. How big is the random walk in GNP? *Journal of Political Economy* 96, 893–920.

Delgado, M.A. 1996. Testing serial independence using the sample distribution function. *Journal of Time Series Analysis* 17, 271–85.

Deo, R.S. 2000. Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *Journal of Econometrics* 99, 291–315.

Durbin, J. 1970. Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables. *Econometrica* 38, 422–1.

Durbin, J. and Watson, G.S. 1950. Testing for serial correlation in least squares regression: I. *Biometrika* 37, 409–28.

Durbin, J. and Watson, G.S. 1951. Testing for serial correlation in least squares regression: II. *Biometrika* 38, 159–78.

Durlauf, S.N. 1991. Spectral based testing of the martingale hypothesis. *Journal of Econometrics* 50, 355–76.

Engle, R. 1982. Autoregressive conditional hetersokedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1008.

Engle, R., Lilien, D. and Robins, R.P. 1987. Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica* 55, 391–407.

Engle, R. and Manganelli, S. 2004. CARViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* 22, 367–91.

Fan, J. and Zhang, W. 2004. Generalized likelihood ratio tests for spectral density. *Biometrika* 91, 195–209.

Glosten, R., Jagannathan, R. and Runkle, D. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–801.

Godfrey, L.G. 1978. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica* 46, 1293–301.

Granger, C.W.J. and Lin, J.L. 1994. Using the mutual information coefficient to identify lags in nonlinear models. *Journal of Time Series Analysis* 15, 371–84.

Granger, C.J.W. and Terasvirta, T. 1993. *Modeling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

Harvey, C.R. and Siddique, A. 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 51, 1263–95.

Hayashi, F. 2000. *Econometrics*. Princeton: Princeton University Press.

Hong, Y. 1996. Consistent testing for serial correlation of unknown form. *Econometrica* 64, 837–64.

Hong, Y. 1998. Testing for pairwise serial independence via the empirical distribution function. *Journal of the Royal Statistical Society, Series B* 60, 429–53.

Hong, Y. 1999. Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *Journal of the*

American Statistical Association 94, 1201–20.

Hong, Y. 2000. Generalized spectral tests for serial dependence. *Journal of the Royal Statistical Society, Series B* 62, 557–74.

Hong, Y. and Lee, T.H. 2003a. Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. *Review of Economics and Statistics* 85, 1048–62.

Hong, Y. and Lee, T.H. 2003b. Diagnostic checking for the adequacy of nonlinear time series models. *Econometric Theory* 19, 1065–121.

Hong, Y. and Lee, Y.J. 2005. Generalized spectral testing for conditional mean models in time series with conditional heteroskedasticity of unknown form. *Review of Economic Studies* 72, 499–51.

Hong, Y. and Lee, Y.J. 2007. Consistent testing for serial correlation of unknown form under general conditional heteroskedasticity. Working paper, Department of Economics, Cornell University, and Department of Economics, Indiana University.

Hong, Y. and White, H. 2005. Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica* 73, 837–901.

Hsieh, D.A. 1989. Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business* 62, 339–68.

Ljung, G.M. and Box, G.E.P. 1978. On a measure of lack of fit in time series models. *Biometrika* 65, 297–303.

Lo, A.W. and MacKinlay, A.C. 1988. Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies* 1, 41–66.

Nelson, D. 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* 59, 347–70.

Newey, W.K. and West, K.D. 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–8.

Paparoditis, E. 2000. Spectral density based goodness-of-fit tests for time series models. *Scandinavian Journal of Statistics* 27, 143–76.

Priestley, M.B. 1988. *Non-Linear and Non-Stationary Time Series Analysis*. London: Academic Press.

Robinson, P.M. 1991. Consistent nonparametric entropy-based testing. *Review of Economic Studies* 58, 437–53.

Robinson, P.M. 1994. Time series with strong dependence. In *Advances in Econometrics, Sixth World Congress*, vol. 1, ed. C. Sims. Cambridge: Cambridge University Press.

Skaug, H.J. and Tjøstheim, D. 1993a. Nonparametric tests of serial independence. In *Developments in Time Series Analysis*, ed. S. Rao. London: Chapman and Hall.

Skaug, H.J. and Tjøstheim, D. 1993b. A nonparametric test of serial independence based on the empirical distribution function. *Biometrika* 80, 591–602.

Skaug, H.J. and Tjøstheim, D. 1996. Measures of distance between densities with application to testing for serial independence. In *Time Series Analysis in Memory of E.J. Hannan*, ed. P. Robinson and M. Rosenblatt. New York: Springer.

Tjøstheim, D. 1996. Measures and tests of independence: a survey. *Statistics* 28, 249–84.

Vinod, H.D. 1973. Generalization of the Durbin–Watson statistic for higher order autoregressive processes. *Communications in Statistics* 2, 115–44.

Wallis, K.F. 1972. Testing for fourth order autocorrelation in quarterly regression equations. *Econometrica* 40, 617–36.

Whang, Y.J. 1998. A test of autocorrelation in the presence of heteroskedasticity of unknown form. *Econometric Theory* 14, 87–122.

Wooldridge, J.M. 1990. An encompassing approach to conditional mean tests with applications to testing nonnested hypotheses. *Journal of Econometrics* 45, 331–50.

Wooldridge, J.M. 1991. On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. *Journal of Econometrics* 47, 5–46.

How to cite this article

Hong, Yongmiao. "serial correlation and serial dependence." The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008. The New Palgrave Dictionary of Economics Online. Palgrave Macmillan. 04 August 2008 <http://www.dictionaryofeconomics.com/article?id=pde2008_S000485> doi:10.1057/9780230226203.1514