since by the very nature of a goal or an end people wish to achieve that goal as soon as possible.

MURRAY N. ROTHBARD

See also **Böhm-Bawerk, Eugen von; Fisher, Irving.**

Bibliography

Bailey, S. 1825. *A Critical Dissertation on the Nature, Measure, and Causes of Value*. New York: Kelley, 1967.

Böhm-Bawerk, E. von. 1884–9. *Capital and Interest*, vols 1 and 2, 4th edn, South Holland, IL: Libertarian Press, 1959.

Fetter, F.A. 1902. The 'roundabout process' in the interest theory. *Quarterly Journal of Economics* 17(November), 163–80. Repr. in Fetter (1977).

Fetter, F.A. 1915. *Economic Principles*, vol. 1. New York: The Century Co.

Fetter, F.A. 1977. *Capital, Interest, and Rent: Essays in the Theory of Distribution*, ed. M. Rothbard. Kansas City: Sheed Andrews and McMeel.

Fisher, I. 1907. *The Rate of Interest*. New York: Macmillan.

Fisher, I. 1930. *The Theory of Interest*. New York: Kelley & Millman, 1954.

Gordon, B. 1975. *Economic Analysis Before Adam Smith: Hesiod to Lessius*. New York: Barnes & Noble.

Kauder, E. 1965. *A History of Marginal Utility Theory*. Princeton: Princeton University Press.

Longfield, S.M. 1971. *The Economic Writings of Mountifort Longfield*, ed. R.D.C. Black. Clifton, NJ: Kelley.

Menger, C. 1871. *Principles of Economics*. Ed. J. Dingwall and B. Hoselitz, Glencoe, IL: Free Press, 1950.

Mises, L. von. 1949. *Human Action: A Treatise on Economics*. 3rd revised edn, Chicago: Regnery, 1966.

Monroe, A., ed. 1924. *Early Economic Thought*. Cambridge, MA: Harvard University Press.

Noonan, J.T., Jr. 1957. *The Scholastic Analysis of Usury*. Cambridge, MA: Harvard University Press.

Rae, J. 1834. Some new principles on the subject of political economy. In *John Rae: Political Economist*, ed. R.W. James. Toronto: University of Toronto Press, 1965.

Turgot, A.R.J. 1977. *The Economics of A.R.J. Turgot*, ed. P.D. Groenewegen. The Hague: Martinus Nijhoff.

Wicksell, K. 1911. Böhm-Bawerk's theory of interest. In K. Wicksell, *Selected Papers on Economic Theory*, ed. E. Lindahl. Cambridge, MA: Harvard University Press, 1958.

Wicksell, K. 1924. The new edition of Menger's *Grundsatze*. In K. Wicksell, *Selected Papers on Economic Theory*, ed. E. Lindahl. Cambridge, MA: Harvard University Press, 1958.

time series analysis

Any series of observations ordered along a single dimension, such as time, may be thought of as a time series. The emphasis in time series analysis is on studying the dependence among observations at different points in time. What distinguishes time series analysis from general multivariate analysis is precisely the temporal order imposed on the observations. Many economic variables, such as GNP and its components, price indices, sales, and stock returns are observed over time. In addition to being interested in the contemporaneous relationships among such variables, we are often concerned with relationships between their current and past values, that is, relationships over time.

The study of time series of, for example, astronomical observations predates recorded history. Early writers on economic subjects occasionally made explicit reference to astronomy as the source of their ideas. For example, Cournot (1838) stressed that, as in astronomy, it is necessary to recognize the *secular* variations which are independent of the periodic variations. Similarly, Jevons (1884) remarked that his study of short-term fluctuations used the methods of astronomy and meteorology. During the 19th century interest in, and analysis of, social and economic time series evolved into a new field of study independent of developments in astronomy and meteorology (see Nerlove, Grether and Carvalho, 1979, pp. 1–21, for a historical survey).

Harmonic analysis is one of the earliest methods of analysing time series thought to exhibit some form of periodicity. In this type of analysis, the time series, or some simple transformation of it, is assumed to be the result of the superposition of sine and cosine waves of different frequencies. However, since summing a finite number of such strictly periodic functions always results in a perfectly periodic series, which is seldom observed in practice, one usually allows for an additive stochastic component, sometimes called 'noise'. Thus, an observer must confront the problem of searching for 'hidden periodicities' in the data, that is, the unknown frequencies and amplitudes of sinusoidal fluctuations hidden amidst noise. An early method for this purpose is *periodogram analysis*, suggested by Stokes (1879) and used by Schuster (1898) to analyse sunspot data and later by others, principally William Beveridge (1921; 1922), to analyse economic time series.

Spectral analysis is a modernized version of periodogram analysis modified to take account of the stochastic nature of the entire time series, not just the noise component. If it is assumed that economic time series are fully stochastic, it follows that the older periodogram technique is inappropriate and that considerable difficulties in the interpretation of the periodograms of economic series may be encountered.

At the time when harmonic analysis proved to be inadequate for the analysis of economic and social time series, another way of characterizing such series was suggested by the Russian statistician and economist, Eugen Slutsky (1927), and by the British statistician, G.U. Yule (1921; 1926; 1927). Slutsky and Yule showed that, if we begin with a series of purely random numbers and then take sums or differences, weighted or unweighted, of

S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*

such numbers, the new series so produced has many of the apparent cyclic properties that were thought at the time to characterize economic and other time series. Such sums or differences of purely random numbers and sums or differences of the resulting series form the basis for the class of autoregressive moving-average (ARMA) processes which are used for modelling many kinds of time series. ARMA models are examples of time domain representations of time series. Although the latter may look very different from spectral representations of time series, there is a one-to-one mapping between time domain analysis and spectral analysis. Which approach is preferred in practice is a matter only of convenience. The choice is often determined by the transparency with which a given question can be answered. The remainder of this article explores these two complementary approaches to the analysis of economic time series.

1. Basic theory

1.1. Stationarity and ergodicity of time series processes

Consider a random variable x_t where $t \in N$, the set of integers; the infinite vector $\{x_t, t \in N\}$ is called a discrete time series. Let M denote a subset of T consecutive elements of N. The distribution of the finite dimensional vector $\{x_t, t \in M\}$ is a well-defined multivariate distribution function, $F_M(\cdot)$. The time series $\{x_t, t \in N\}$ is said to be *strictly stationary* if, for any finite subset M of N and any integer τ, the distribution function of $\{x_t, t \in M + \tau\}$ is the same as the distribution function of $\{x_t, t \in M\}$. In other words, the joint distribution function of the finite vector of observations on x_t is invariant with respect to the origin from which time is measured. All the unconditional moments of the distribution function, if they exist, are independent of the index t; in particular,

$$\begin{aligned} E(x_t) &= \mu \\ \gamma(\tau) &= E[x_t - \mu][x_{t+\tau} - \mu], \end{aligned} \tag{1}$$

where $\gamma(\tau)$ is the autocovariance function and depends only on the difference in indices, τ. Time-series processes for which (1) holds, but which are not necessarily strictly stationary according to the definition above, are said to be weakly *stationary, covariance stationary,* or *stationary to the second order.* Time-series processes for which $F_M(\cdot)$ is multivariate normal for any subset M of N are called *Gaussian processes.* For Gaussian processes covariance stationarity implies strict stationarity.

In practice, we usually observe only one realization of a finite subset of the time series of interest, corresponding to one of the many possible draws of length T from $F_M(\cdot)$. The question is whether the moments of x_t may be inferred from one such realization; for example, from the time averages of sums (or sums of products) of the observed values of a time series. If the process is what is known as *ergodic,* time averages of functions of the observations on the time series at T time points converge in mean square to the corresponding population expectations of x_t across alternative draws, as $T \to \infty$ (Priestley, 1981, pp. 340–3; Doob, 1953, p. 465). It is possible for a process to be stationary, yet not ergodic. Consider, for example, the process $x_t^{(i)} = \eta^{(i)} + \varepsilon_t$, where $x_t^{(i)}$ denotes the *ith* draw for observation x_t from the universe of all possible draws for x_t. Suppose that $\eta^{(i)} \sim N(0, \lambda^2)$ is the mean of the *ith* draw and that $\varepsilon_t \sim N(0, \sigma^2)$ is independent of $\eta^{(i)}$. This process is clearly stationary in that the probability limit of the ensemble average is zero, yet the time average $\sum_{t=1}^{T} x_t^{(i)}/T = \eta^{(i)} + \sum_{t=1}^{T} \varepsilon_t/T$ converges to $\eta^{(i)}$ rather than zero, thus violating ergodicity.

1.2. The Wold decomposition and general linear processes

Let $\{\varepsilon_t\}$ be one element of a time series of serially uncorrelated, identically distributed random variables with zero mean and variance σ^2. Then the infinite, one-sided moving average (MA) process

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \tag{2}$$

where $b_0 = 1$ and $\sum_{j=0}^{\infty} b_j^2 < \infty$, is also a well-defined stationary process with mean 0 and variance $\sigma^2 \Sigma_0^{\infty} b_j^2$. Processes of this form and, more generally, processes based on an infinite two-sided MA of the same form are called *linear processes,* are always ergodic, and play a key role in time series analysis (Hannan, 1970).

The importance of the process (2) is underscored by the Wold decomposition theorem (Wold, 1938), which states that any weakly stationary process may be decomposed into two mutually uncorrelated component processes, one an infinite one-sided MA of the form (2) and the other a so-called linearly deterministic process, future values of which can be predicted exactly by some linear function of past observations. The linearly deterministic component is non-ergodic.

2. Linear processes in time and frequency domains

2.1. Autocovariance and autocovariance generating functions

The autocovariance function of a stationary process, defined in (1) above, or its matrix generalization for vector processes, provides the basic representation of time dependence for weakly stationary processes. For the stationary process defined in (2), it is

$$\gamma(\tau) = \sigma^2 \sum_{j=0}^{\infty} b_j b_{j+\tau}. \tag{3}$$

Let z denote a complex scalar. Then the autocovariance generating transform is defined as

$$g(z) = \sum_{-\infty}^{\infty} \gamma(\tau) z^{\tau} \tag{4}$$

in whatever region of the complex plane the series on the right-hand side converges. If the series $\{x_t\}$ is covariance stationary, convergence will occur in an annulus about the unit circle. The autocovariance generating transform for the one-sided MA process defined in (2) is

$$g(z) = \sigma^2 B(z)B(z^{-1}) \tag{5}$$

where

$$B(z) = \sum_{k=0}^{\infty} b_k z^k.$$

If $B(z)$ has no zeros on the unit circle, the process defined in (2) is invertible and also has an infinite-order autoregressive (AR) representation as

$$A(L)x_t = \varepsilon_t, \tag{6}$$

where L is the lag operator such that $L^j x_t = x_{t-j}$ and $A(L) = a_0 + a_1 L + a_2 L^2 + \ldots$.

So-called ARMA processes have an autocovariance generating transform which is a rational function of z. If the ARMA process is both stationary and invertible, $g(z)$ may be written as

$$G(z) = \frac{P(z)P(z^{-1})}{Q(z)Q(z^{-1})} = \sigma^2 \frac{\prod_{k=1}^{m}(1-\beta_k z)(1-\beta_k z^{-1})}{\prod_{j=1}^{n}(1-\alpha_j z)(1-\alpha_j z^{-1})} \tag{7}$$

where $|\beta_k|, |\alpha_j| < 1 \ \forall j, k$. Then the corresponding ARMA model is

$$Q(L)x_t = P(L)\varepsilon_t, \tag{8}$$

where

$$Q(L) = \prod_{j=1}^{n}(1-\alpha_j L) \text{and} P(L) = \prod_{k=1}^{m}(1-\beta_k L).$$

2.2. Spectral density functions

If the value of z lies on the complex unit circle, it follows that $z = e^{-i\lambda}$, where $i = \sqrt{-1}$ and $-\pi \leq \lambda \leq \pi$. Substituting for z in the autocovariance generating transform (5) and dividing by 2π, we obtain the *spectral density function* of a linearly non-deterministic stationary process $\{x_t\}$ in terms of the frequency λ:

$$\begin{aligned} f(\lambda) &= (1/2\pi)g(e^{i\lambda}) = (\sigma^2/2\pi)B(e^{i\lambda})B(e^{-i\lambda}) \\ &= (1/2\pi)\sum_{-\infty}^{\infty}\gamma(\tau)e^{-i\lambda\tau}, \qquad -\pi \leq \lambda < \pi. \end{aligned} \tag{9}$$

Thus, the spectral density function is the Fourier transform of the autocovariance function. It can be shown that for a process with absolutely summable autocovariances the spectral density function exists and can be used to compute all of the autocovariances, so the same time series can be characterized equivalently in terms of the autocovariance function in the time-domain or in terms of the spectral density function in the frequency domain.

The spectral density function for a linearly non-deterministic, stationary, real-valued time series is a real-valued, non-negative function, symmetric about the origin, defined in the interval $[-\pi, \pi]$:

$$f(\lambda) = (1/2\pi)\left[\gamma(0) + 2\sum_{\tau=1}^{\infty}\gamma(\tau)\cos\lambda\tau\right]. \tag{10}$$

Moreover,

$$E(x_t - \mu)^2 = \int_{-\pi}^{\pi} f(\lambda)\mathrm{d}\lambda, \tag{11}$$

so that the spectral density function is a frequency-band decomposition of the variance of $\{x_t\}$.

When the process generating $\{x_t\}$ is merely stationary, that is, when $\{x_t\}$ may have a linearly deterministic component, the spectral density function is

$$f(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda\tau}\mathrm{d}F(\lambda), \tag{12}$$

where $f(\lambda)$ is a distribution function (Doob, 1953, p. 488). Note that deterministic seasonal effects, for example, may cause a jump in the spectral distribution function.

The autocovariance function, its generating transform and the spectral distribution function all have natural generalizations to the multivariate case, in which $\{x_t\}$ can be thought of as a vector of time-series processes.

The estimation and analysis of spectral density and distribution functions play an important role in all forms of time-series analysis. More detailed treatments are Doob (1953), Fishman (1969), Koopmans (1974), Fuller (1976), Nerlove, Grether and Carvalho (1979, ch. 3) and Priestley (1981).

2.3. Unobserved components (UC) models

In the statistical literature dealing with the analysis of economic time series it is common practice to classify the types of movements that characterize a time series as trend, cyclical, seasonal, and irregular components. The idea that a time series may best be viewed as being composed of several unobserved components is by no means universal, but it plays a fundamental role in many applications, for example, the choice of methods for seasonal adjustment. Nerlove, Grether and Carvalho (1979, ch. 1) review the history of the idea of unobserved components in economics from its origin early in the 19th century.

In the 1960s, Nerlove (1964; 1965; 1967) and Granger (1966) suggested that the typical spectral shape of many economic time series could be accounted for by the superposition of two or more independent components with specified properties. There are basically two approaches to the formulation of UC models: Theil and Wage (1964) and Nerlove and Wage (1964), Nerlove (1967) and Grether and Nerlove (1970) choose the form of components in such a way as to replicate the typical spectral shape of the series which represents their superposition. For example, let T_t represent the trend component, C_t the cyclical, S_t the seasonal, and I_t the irregular of a monthly time series; then the observed series can be represented as

$$y_t = T_t + C_t + S_t + I_t, \tag{13}$$

where

$$T_t = a_0 + a_1 t + a_2 t^2 + \ldots + a_p t^p,$$
$$C_t = \frac{1 + \beta_1 L + \beta_2 L^2}{(1 - \alpha_1 L)(1 - \alpha_2 L)} \varepsilon_{1t},$$
$$S_t = \frac{1 + \beta_3 L + \beta_4 L^2}{1 - \gamma L^{12}} \varepsilon_{2t},$$
$$I_t = \varepsilon_{3t},$$

and $\varepsilon_{1t}, \varepsilon_{2t}$, and ε_{3t} are i.i.d. normal variables with variances σ_{11}, σ_{22}, and σ_{33}, respectively. This approach has been carried forward by Harvey (1984), Harvey and Peters (1990) and Harvey and Todd (1984).

An alternative approach is to derive the components of the UC model from a well-fitting ARMA model (obtained after suitably transforming the data), given sufficient a priori identifying restrictions on the spectral properties of the components. See Box, Hillmer and Tiao (1978), Pierce (1978; 1979), Burman (1980), Hillmer and Tiao (1982), Hillmer, Bell and Tiao (1983), Bell and Hillmer (1984), Burridge and Wallis (1985), and Maravall (1981; 1984). The basis of this procedure is the fact that every stationary UC model, or the stationary part of every UC model, has an equivalent ARMA form, the so-called canonical form of the UC model (Nerlove and Wage, 1964; Nerlove, Grether and Carvalho, 1979, ch. 4).

3. Specification, estimation, inference and prediction

3.1. Autocovariance and spectral density functions

Suppose we have a finite number of observations of a realization of the process generating the time series, say $x_1, \ldots, x_T$. For expository purposes it is assumed that all deterministic components of x_t have been removed. If μ is unknown, this may be accomplished by subtracting the sample mean of the time series observations from the data prior to the analysis. For a zero mean series x_t there are basically two ways of estimating $\gamma(\tau)$ defined in (1): the first is the biased estimator

$$c(\tau) = (1/T) \sum_{t=1}^{T-|\tau|} x_t x_{t+|\tau|},$$
$$\tau = 0, \pm 1, \ldots, \pm M, M \leq (T-1). \tag{14}$$

The second is the unbiased estimator

$$\tilde{c}(\tau) = [1/(T - |\tau|)] \sum_{t=1}^{T-|\tau|} x_t x_{t+|\tau|},$$
$$\tau = 0, \pm 1, \ldots, \pm M, \quad M \leq T - 1. \tag{15}$$

Although $c(\tau)$ is biased in finite samples, it is asymptotically unbiased. The key difference between $c(\tau)$ and $\tilde{c}(\tau)$ is that $c(\tau)$ is a positive definite function of τ whereas $\tilde{c}(\tau)$ is not (Parzen, 1961, p. 981). The variance and covariances of the estimated autocovariances are derived, *inter alia*, by Hannan (1960), and Anderson (1971). As $T \to \infty$, both tend to zero, as the estimates are asymptotically uncorrelated and consistent. However,

$$E[c(\tau) - Ec(\tau)]^2 / E[c(\tau)] \to \infty \quad \text{as} \quad \tau/T \to 1. \tag{16}$$

This property accounts for the failure of the estimated autocorrelation function

$$r(\tau) = c(\tau)/c(0) \tag{17}$$

to damp down as $\tau \to \infty$, as it should for a stationary, linearly non-deterministic process (Hannan, 1960, p. 43).

A 'natural' estimator of the spectral density function is obtained by replacing $\gamma(\tau)$ in (10) by $c(\tau)$ or $\tilde{c}(\tau)$. The resulting estimator is proportional, at each frequency, to a sample quantity called the periodogram:

$$I_T(\lambda) = (2/T) \left| \sum_{1}^{T} e^{i\lambda t} x_t \right|^2 \tag{18}$$

usually evaluated at the equi-spaced frequencies

$$\lambda = 2k\pi/T, \quad k = 1, 2, \ldots, [T/2] \tag{19}$$

in the interval $[0, \pi]$. Although, for a stationary, non-linearly deterministic process, the periodogram ordinates are asymptotically unbiased estimates of the spectral densities at the corresponding frequencies, they are not consistent estimates; moreover, the correlation between adjacent periodogram ordinates tends to zero with increasing sample size. The result is that the periodogram presents a jagged appearance which is increasingly difficult to interpret as more data become available.

In order to obtain consistent estimates of the spectral density function at specific frequencies, it is common practice to weight the periodogram ordinates over the frequency range or to form weighted averages of the autocovariances at different lags. There is a substantial literature on the subject. The weights are called a 'spectral window'. Essentially the idea is to reduce the variance of the estimate of an average spectral density around a particular frequency by averaging periodogram ordinates which are asymptotically unbiased and independently distributed estimates of the corresponding ordinates of the spectral density function. Related weights can also be applied to the estimated autocovariances which are substituted in (10); this weighting system is called a 'lag window'.

Naturally the sampling properties of the spectral estimates depend on the nature of the 'window' used to obtain consistency (see Priestley, 1981, pp. 432–94 for further discussion). Regardless of the choice of window, the 'bandwidth' used in constructing the window must decrease at a suitable rate as the sample size grows. In the spectral window approach, this means that the window width must decrease at a slower rate than the sample size. In the lag window approach, this means that the number of included autocovariances must increase at a slower rate than the sample size.

3.2. ARMA models

The autocovariance function and the spectral density function for a time series represent nonparametric approaches to describing the data. An alternative approach is to specify and estimate a parametric ARMA model for x_t This approach involves choosing the orders of the polynomials P and Q in (7) and (8) and perhaps also specifying that one or more coefficients are zero or placing other restrictions on P and Q. The problem then becomes one of estimating the parameters of the model.

Despite the poor statistical properties of the estimated autocovariance function and a related function called the partial autocorrelation function, these are sometimes used to specify the orders of the polynomials P and Q. An alternative approach is to select the model that minimizes the value of information-theoretic criteria of the form

$$\mathrm{IC}(i) = \log(\hat{\sigma}_i^2) + k_i c_T, \qquad (20)$$

where k_i refers to the number of estimated parameters in the candidate models $i = 1, \ldots, M$, and $\hat{\sigma}_i^2$ to the corresponding maximum likelihood estimate of the residual variance. Such criteria incorporate a trade-off between the fit of a model and its degree of parsimony. That trade-off depends on the penalty term c_t (Akaike, 1970; 1974; Schwarz, 1978). There is no universally accepted choice for c_T. For $c_T = 2/T$ expression (20) reduces to the Akaike information criterion (AIC), for example, and for $c_T = \ln(T)/T$ to the Schwarz information criterion (SIC). The asymptotic properties of alternative criteria will depend on the objective of the user and the class of models considered.

Given the orders of the AR and MA components, a variety of maximum likelihood or approximate maximum likelihood methods are available to estimate the model parameters. Newbold (1974) shows that, if x_t is characterized by (8) with $\varepsilon_t \sim NID(0, \sigma^2)$, then the exact likelihood function for the parameters of $P(\cdot)$ and $Q(\cdot)$ is such that the maximum likelihood estimates of the parameters and the least-squares (LS) estimates (in general highly nonlinear) are asymptotically identical. Only in the case of a pure AR model are the estimates linear conditional on the initial observations. Several approximations have been discussed (Box and Jenkins, 1970; Granger and Newbold, 1977; Nerlove, Grether and Carvalho, 1979, pp. 121–5).

Exact maximum likelihood estimation of ARMA models has been discussed by, inter alia, Newbold (1974), Anderson (1977), Ansley (1979), and Harvey (1981). Following Schweppe (1965), Harvey suggests the use of the Kalman filter to obtain the value of the exact-likelihood function, which may be maximized by numerical methods. The Kalman filter approach is easily adapted to the estimation of UC models in the time domain.

An alternative to exact or approximate maximum-likelihood estimation in the time domain was suggested by Hannan (1969). Estimates may be obtained by maximizing an approximate likelihood function based on the asymptotic distribution of the periodogram ordinates defined in (18). These are asymptotically independently distributed (Brillinger, 1975, p. 95), and the random variables $2I_t(\lambda)/f(\lambda)$ have an asymptotic χ^2 distribution with two degrees of freedom (Koopmans, 1974, pp. 260–5). This means that the asymptotic distribution of the observations, $\{x_1, \ldots, x_T\}$ is proportional to

$$\prod_{j=0}^{[T/2]} \left[1/f(\lambda_j)\right] \exp\left[-I(\lambda_j)/f(\lambda_j)\right] \qquad (21)$$

where $\lambda_j = 2j\pi/T, j = 0, \ldots, [T/2]$, are the equi-spaced frequencies in the interval $[0, \pi]$ at which the periodogram is evaluated (Nerlove, Grether and Carvalho, 1979, pp. 132–6). Since the true spectral density $f(\lambda)$ depends on the parameters characterizing the process, this asymptotic distribution may be interpreted as a likelihood function. Frequency domain methods, as these are called, may easily be applied in the case of UC models.

Whether approximate or exact maximum-likelihood estimation methods are employed, inference may be based on the usual criteria related to the likelihood function. Unfortunately, serious difficulties may be encountered in applying the asymptotic theory, since the small sample distribution of the maximum likelihood estimator may differ greatly from the limiting distribution in important cases (Sargan and Bhargava, 1983; Anderson and Takemura, 1986).

3.3. Prediction and extraction

The problem of prediction is essentially the estimation of an unknown future value of the time series itself; the problem of extraction, best viewed in the context of UC models described in section 2.3, is to estimate the value of one of the unobserved components at a particular point in time, not necessarily in the future. Problems of trend extraction and seasonal adjustment may be viewed in this way (Grether and Nerlove, 1970). How the prediction (or extraction) problem is approached depends on whether we are assumed to have an infinite past history and, if not, whether the parameters of the process generating the time series are assumed to be known. In practice, of course, an infinite past history is never available, but a very long history is nearly equivalent if the process is stationary or can be transformed to stationarity. It is common, as well, to restrict attention to linear predictors, which involves no loss of generality if the processes considered are Gaussian and little loss if merely linear. To devise a theory of optimal prediction or extraction requires some criterion by which to measure the accuracy of a particular candidate. The most common choice is the minimum mean-square error (MMSE) criterion, which is also the conditional expectation of the unknown quantity. For a discussion of alternative loss functions see Granger (1969) and Christoffersen and Diebold (1996; 1997).

The theory of optimal prediction and extraction due to Kolmogorov (1941) and Wiener (1949) and elaborated by Whittle (1963) for discrete processes assumes a possibly infinite past history and known parameters. As a special case of the Wiener–Kolmogorov theory for nondeterministic, stationary processes, consider the linear process defined by (2). Since the ε_t are i.i.d. with zero mean and variance σ^2, it is apparent that the conditional expectation of x_{t+v}, given all innovations from the infinite past to t, is

$$\hat{x}_{t+v} = b_v\varepsilon_t + b_{v+1}\varepsilon_{t+1} + \ldots \qquad (22)$$

Of course, even if the parameters $b_j, j = 0, 1, \ldots,$ are assumed to be known, the series $\{\varepsilon_t\}$ is not directly observable. The ε_t's are sometimes called the *innovations* of the process, since it is easy to show that $\varepsilon_{t+1} = x_{t+1} - \hat{x}_{t+1}$ is the one-step ahead prediction error. If the process is invertible, it has the autoregressive representation (6) and so can be expressed solely in terms of the, generally infinite-order, autoregression

$$\hat{x}_{t+v} = D(L)x_t, \qquad (23)$$

where the generating transform of the coefficients of D is

$$D(z) = \frac{1}{B(z)}\left[\frac{B(z)}{z^v}\right]_+.$$

The operator $[\cdot]_+$ eliminates terms involving negative powers of z.

The problem of extraction is best viewed in the context of multiple time series; in general we wish to 'predict' one time series $\{y_t\}$ from another related series $\{x_t\}$. It is not necessary that the series $\{y_t\}$ actually be observed as long as its relationship to an observed series $\{x_t\}$ can be described (Nerlove, Grether and Carvalho, 1979, ch. 5).

The Kalman filter approach to prediction and extraction (Kalman, 1960) is both more special and more general than the Wiener–Kolmogorov theory: attention is restricted to finite-dimensional parameter spaces and linear processes, but these processes need not be stationary. The parameters may vary with time, and we do not require an infinite past. This approach represents a powerful tool of practical time-series analysis and may be easily extended to multiple time series. A full discussion, however, requires a discussion of 'state-space representation' of time series processes and is beyond the scope of this entry (Harvey, 1989).

4. Multiple time series analysis

A general treatment of multiple time series analysis is contained in Hannan (1970). The two-variable case will serve to illustrate the matter. Two stationary time series $\{x_t\}$ and $\{y_t\}$ are said to be jointly stationary if their joint distribution function does not depend on the origin from which time is measured. Joint stationarity implies, but is not in general implied by, weak or covariance joint stationarity; that is, $cov(x_t, y_s)$ is a function of $s - t$ only. In this case the cross-covariance function is

$$\gamma_{yx}(\tau) = E\left[y_t - \mu_y\right]\left[x_{t-\tau} - \mu_x\right], \qquad (24)$$

where $\mu_x = Ex_t$ and $\mu_y = Ey_t$. Note that $\gamma_{yx}(\tau)$ and $\gamma_{xy}(\tau)$ are, in general, different. The cross-covariance generating function is defined as

$$g_{yx}(z) = \sum_{-\infty}^{\infty}\gamma_{yx}(\tau)z^{\tau} \qquad (25)$$

in that region of the complex plane in which the right-hand side of (25) converges. For two jointly stationary series this occurs in an annulus containing the unit circle. In this case, the cross-spectral density function is defined as

$$f_{yx}(\lambda) = (1/2\pi)g_{yx}\left(e^{i\lambda}\right). \qquad (26)$$

Since $\gamma_{yx}(\tau)$ and $\gamma_{xy}(\tau)$ are not equal, the cross-spectral density function is complex valued and can be decomposed into a real part (the co-spectral density) and a complex part (the quadrature spectral density):

$$f_{yx}(\lambda) = c_{yx}(\lambda) + iq_{yx}(\lambda). \qquad (27)$$

In polar form, the cross-spectral density may be written as

$$f_{yx}(\lambda) = \alpha_{yx}(\lambda)\exp\left[i\phi_{yx}(\lambda)\right], \qquad (28)$$

where $\alpha_{yx}(\lambda) = [c^2_{yx}(\lambda) + q^2_{yx}(\lambda)]^{1/2}$ is called the amplitude or gain, and where $\phi_{yx}(\lambda) = \arctan\{-q_{yx}(\lambda)/c_{yx}(\lambda)\}$ is called the phase. Another useful magnitude is the coherence between the two series, defined as

$$\rho_{yx}(\lambda) = \frac{\left|f_{yx}(\lambda)\right|^2}{f_{xx}(\lambda)f_{yy}(\lambda)}, \qquad (29)$$

which measures the squared correlation between y and x at a frequency λ. Clearly, $\rho_{yx}(\lambda) = \rho_{xy}(\lambda)$. Estimation of cross-spectral density functions and related quantities is discussed in Priestley (1981, pp. 692–712).

Often it is convenient to impose additional parametric structure in modelling multiple time series. The workhorse multiple time series model in econometrics has been the covariance-stationary K-dimensional vector autoregressive model, which may be viewed as a natural generalization of the univariate AR model discussed earlier:

$$A(L)x_t = \varepsilon_t \qquad (30)$$

where $A(L) = I_K - A_1L - \ldots - A_pL^p$. Here each variable in x_t is regressed on its own lags as well as lags of all other variables in x_t up to some pre-specified lag order p. This *vector autoregression* (VAR) can also be viewed as an approximation to a general linear process x_t, and may be estimated by LS.

Similarly, the formulation of ARMA and UC models discussed earlier may be extended to the multivariate case by interpreting the polynomials in the lag operator as matrix polynomials and by replacing the scalar random variables by vectors. Although these vector ARMA and UC models bear a superficial resemblance to the corresponding univariate ones, their structure is, in fact, much more complicated and gives rise to difficult identification problems. In the univariate case, we can formulate simple conditions under which a given covariance function identifies a unique ARMA or UC model, but in the multivariate case these conditions are no longer sufficient. Hannan (1970; 1971) gives a complete treatment. State-space methods have also been employed to study the structure of multivariate ARMA models (Hannan, 1976; and, especially, 1979).

5. Unit roots, co-integration and long memory

Standard tools for time series analysis have been developed for processes that are covariance stationary or have been suitably transformed to achieve covariance stationarity by removing (or explicitly modelling) deterministic trends, structural breaks, and seasonal effects. The presence of a *unit root* in the autoregressive lag order polynomial of an ARMA process also violates the assumption of stationarity. Processes with a unit root are also called integrated of order one (or $I(1)$ for short) because they become covariance-stationary only upon being differenced once. In general, $I(d)$ processes must be differenced d times to render the process covariance-stationary.

The presence of unit roots has important implications for estimation and inference. When the scalar process x_t is $I(1)$ the variance of x_t will be unbounded, model innovations will have permanent effects on the level of x_t, the autocorrelation function does not die out, and x_t will not revert to a long-run mean. Moreover, coefficients of $I(1)$ regressors will have nonstandard asymptotic distributions, invalidating standard tools of inference.

The simplest example of an autoregressive integrated moving-average (ARIMA) process is the random walk process: $x_t = x_{t-1} + \varepsilon_t$. The potential pitfalls of regression analysis with $I(1)$ data are best illustrated by the problem of regressing one independent random walk on another. In that case, it can be shown that R^2 and $\hat{\beta}$ will be random and that the usual t-statistic will diverge, giving rise to seemingly significant correlations between variables that are unrelated by construction. This *spurious regression* problem was first discussed by Yule (1926), further illustrated by Granger and Newbold (1974), and formally analyzed by Phillips (1986) and Phillips and Durlauf (1986). Similar problems arise in deterministically detrending $I(1)$ series (Nelson and Kang, 1981; Durlauf and Phillips, 1988). Unbalanced regressions, that is, regressions in which the regressand is not of the same order of integration as the regressor, may also result in spurious inference. An exception to this rule is inference on coefficients of mean zero $I(0)$ variables in regressions that include a constant term (Sims, Stock and Watson, 1991).

The standard response to dealing with $I(1)$ data is to difference the data prior to the analysis. There is one important exception to this rule. There are situations in which several variables are individually $I(1)$, but share a common unit root component. In that case, a linear combination of these variables will be $I(0)$:

$$c'x_t = u_t \sim I(0), \quad c \neq 0 \qquad (31)$$

where x_t denotes a K-dimensional vector of $I(1)$ variables and c is a $(K \times 1)$ parameter vector. In other words, these variables share a common stochastic trend. This phenomenon is known as *co-integration* (Granger, 1981; Engle and Granger, 1987) and c is known as the co-integrating vector. Clearly, c is not unique. It is common to normalize one element of c to unity. The LS estimator of c in (31) is consistent, but corrections for omitted dynamics are recommended (Stock and Watson, 1993; Phillips and Hansen, 1990). Co-integrating relationships have been used extensively in modelling long-run equilibrium relationships in economic data (Engle and Granger, 1991).

Variables that are co-integrated are linked by an *error correction* mechanism that prevents the integrated variables from drifting apart without bound. Specifically, by

the Granger representation theorem of Engle and Granger (1987), under some regularity conditions, any K-dimensional vector of co-integrated variables x_t can be represented as a vector error correction (VEC) model of the form:

$$\Delta x_t = \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} - \Pi x_{t-p} \qquad (32)$$

where Γ_i, $i = 1, \ldots, p-1$, and $\Pi \equiv BC$ are conformable coefficient matrices and Δ denotes the first-difference operator. Model (32) allows for up to r co-integrating relationships where r is the rank of Π. For $r = 0$, the error correction term in model (32) drops out and the model reduces to a difference-stationary VAR. For $r = K$, all variables are $I(0)$ and model (32) is equivalent to a stationary VAR in levels. Otherwise, there are $0 < r < K$ common trends. If the $(r \times K)$ matrix of co-integrating vectors, C, is known, the model in (32) reduces to

$$\Delta x_t = \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} - B z_{t-p} \qquad (32')$$

where $z_{t-p} \equiv C x_{t-p}$, and the model may be estimated by LS; if only the rank r is known, the VEC model in (32) is commonly estimated by full information maximum likelihood methods (Johansen, 1995).

Starting with Nelson and Plosser (1982), a large literature has dealt with the problem of statistically discriminating between $I(1)$ and $I(0)$ models for economic data. Notwithstanding these efforts, it has remained difficult to detect reliably the existence of a unit root (or of co-integration). The problem is that in small samples highly persistent, yet stationary processes are observationally equivalent to exact unit root processes. It may seem that not much could hinge on this distinction then, but it can be shown that $I(1)$ and $I(0)$ specifications that fit the data about equally well may have very different statistical properties and economic implications (Rudebusch, 1993).

For processes with roots near unity in many cases neither the traditional asymptotic theory for $I(0)$ processes nor the alternative asymptotic theory for exact $I(1)$ processes will provide a good small-sample approximation to the distribution of estimators and test statistics. An alternative approach is to model the dominant root, ρ, of the autoregressive lag order polynomial as *local-to-unity* in the sense that $\rho = 1 - c/T$, $c > 0$. This asymptotic thought experiment gives rise to an alternative asymptotic approximation that in many cases provides a better small-sample approximation than imposing the order of integration or relying on unit root pretests (Stock, 1991; Elliott, 1998).

Stationary ARMA processes are 'short memory' processes in that their autocorrelation function dies out quickly. For large τ, ARMA autocorrelations decay approximately geometrically, that is, $\rho(\tau) \approx r^\tau$, where r is a constant such that $|r| < 1$. In many applied contexts including volatility dynamics in asset returns, there is evidence that the autocorrelation function dies out much more slowly. This observation has motivated the development of the class of *fractionally integrated* ARMA (ARFIMA) models:

$$Q(L)(1 - L)^d x_t = P(L)\varepsilon_t \qquad (33)$$

where d is a real number, as opposed to an integer (Baillie, 1996). Stationarity and invertibility require $|d| < 0.5$, which can always be achieved by taking a suitable number of differences. The autocorrelation function of an ARFIMA process decays at a hyperbolic rate. For large τ, we have $\rho(\tau) \approx \tau^{2d-1}$, where $d < 1/2$ and $d \neq 0$. Such '*long memory*' models may be estimated by the two-step procedure of Geweke and Porter-Hudak (1983) or by maximum likelihood (Sowell, 1992; Baillie, Bollerslev and Mikkelsen, 1996). A detailed discussion including extensions to the notion of fractional co-integration is provided by Baillie (1996). Long memory may arise, for example, from infrequent stochastic regime changes (Diebold and Inoue, 2001) or from the aggregation of economic data (Granger, 1980; Chambers, 1998). Perhaps the most successful application of long-memory processes in economics has been work on modelling the volatility of asset prices and powers of asset returns, yielding new insights into the behaviour of markets and the pricing of financial risk.

6. Nonlinear time series models

The behaviour of many economic time series appears to change distinctly at irregular intervals, consistent with economic models that suggest the existence of floors and ceilings, buffer stocks and regime switches in the data. This observation has given rise to a large literature dealing with nonlinear time series models. Nonlinear time series models still have a Wold representation with linearly unpredictable innovations, but these innovations are nevertheless dependent over time. This has important implications for forecasting and for the dynamic properties of the model. For example, the effects of innovations in nonlinear models will depend on the path of the time series and the size of the innovation, and may be asymmetric.

6.1. Nonlinear dynamics in the conditional mean

The increasing importance of nonlinear time series models in econometrics is best illustrated by two examples: hidden Markov chain models and smooth transition regression models of the conditional mean.

The idea of hidden Markov chains first attracted attention in econometrics in the context of *regime switching models* (Hamilton, 1989). The original motivation was that many economic time series appear to

follow a different process during recession phases of the business cycle than during economic expansions. This type of regime-switching behaviour may be modelled in terms of an unobserved discrete-valued state variable (for example, 1 for a recession and 0 for an expansion) that is driven by a Markov chain. The transition from one state to another is governed by a matrix of transition probabilities that may be estimated from past data. The essence of this method thus is that the future will in some sense be like the past. A simple example of this idea is the regime-switching AR(1) model:

$$x_t = a_{1st}x_{t-1} + \varepsilon_t, \varepsilon_t \sim NID(0, \sigma^2) \qquad (34)$$

where the regime s_t is the outcome of an unobserved two-state Markov chain with s_t independent of ε_τ for all t and τ. In this model, the time-varying slope parameter will take on different values depending on the state s. Once the model has been estimated by maximum likelihood methods, it is possible to infer how likely a given regime is to have generated the observed data at date t. An excellent review of the literature on hidden Markov models is provided by Cappé, Moulines and Ryden (2005); for a general treatment of state space representations of nonlinear models, also see Durbin and Koopman (2001).

The idea of *smooth transition regression models* is based on the observation that many economic variables are sluggish and will not move until some state variable exceeds a certain threshold. For example, price arbitrage in markets will only set in once the expected profit of a trade exceeds the transaction cost. This observation has led to the development of models with fixed thresholds that depend on some observable state variable. Smooth transition models allow for the possibility that this transition occurs not all of a sudden at a fixed threshold but gradually, as one would expect in time series data that have been aggregated across many market participants. A simple example is the smooth-transition AR(1) model:

$$\begin{aligned} x_t &= \Phi(z_{t-1}, ..., z_{t-d}, \Gamma)x_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim NID(0, \sigma^2) \end{aligned} \qquad (35)$$

where $\Phi(.)$ denotes the transition function, z_t is a zero mean state variable denoting the current deviation of x_t from a (possibly time-varying) equilibrium level and Γ is the vector of transition parameters. Common choices for the transition function are the logistic or the exponential function. For example, we may specify $\Phi(.) = (\exp\{\gamma(z_{t-1})^2\})$ with $\gamma < 0$. If $z_{t-1} = 0$, $\Phi(.) = 1$ and the model in (35) reduces to a random walk model; otherwise, $\Phi(.) < 1$ and the model in (35) reduces to a stationary AR(1). The degree of mean reversion is increasing in the deviation from equilibrium. For further discussion see Granger and Teräsvirta (1993).

6.2. Nonlinear dynamics in the conditional variance

While the preceding examples focused on nonlinear dynamics in the conditional mean, nonlinearities may also arise in higher moments. The leading example is the conditional variance. Many economic and financial time series are characterized by volatility clustering. Often interest centres on predicting these *volatility dynamics* rather than the conditional mean. The basic idea of modelling and forecasting volatility was set out in Engle's (1982) path-breaking paper on autoregressive conditional heteroskedasticity (ARCH). Subsequently, Bollerslev (1986) introduced the class of generalized autoregressive conditionally heteroskedastic (GARCH). Consider a decomposition of x_t into the one-step ahead conditional mean, $\mu_{t|t-1} \equiv E(x_t|\Omega_{t-1})$, and conditional variance, $\sigma^2_{t|t-1} \equiv Var(x_t|\Omega_{t-1})$, where Ω_{t-1} denotes the information set at $t-1$:

$$x_t = \mu_{t|t-1} + \sigma_{t|t-1}v_t \qquad v_t \sim NID(0, 1) \qquad (36)$$

The leading example of a GARCH model of the conditional variance is the GARCH(1,1) model, which is defined by the recursive relationship

$$\sigma^2_{t|t-1} = \omega + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2} \qquad (37)$$

where $\varepsilon_t \equiv \sigma_{t|t-1}v_t$, and the parameter restrictions $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ ensure that the conditional variance remains positive for all realizations of v_t. The standard estimation method is maximum likelihood. The basic GARCH(1,1) model may be extended to include higher-order lags, to allow the distribution of v_t to have fat tails, to allow for asymmetries in the volatility dynamics, to permit the conditional variance to affect the conditional mean, and to allow volatility shocks to have permanent effects or volatility to have long memory. It may also be extended to the multivariate case.

It follows directly from the formulation of the GARCH(1,1) model that the optimal, in the MMSE sense, one-step-ahead forecast equals $\sigma^2_{t+1|t}$. Similar expressions for longer horizons may be obtained by recursive updating. There is a direct link from the arrival of news to volatility measures and from volatility forecasts to risk assessments. These and alternative volatility models and the uses of volatility forecasts are surveyed in Andersen et al. (2006). For a comparison of GARCH models with the related and complementary class of stochastic volatility models, see Andersen, Bollerslev and Diebold (2006) and Shephard (2005).

7. Applications

Time series analytic methods have many applications in economics. Here we consider five: (1) analysis of the cyclic properties of economic time series, (2) description of seasonality and seasonal adjustment, (3) forecasting,

(4) dynamic econometric modelling, and (5) structural vector autoregressions.

7.1. Analysis of the cyclic properties of economic time series

Suppose that the time series $\{x_t\}$ is a linearly non-deterministic stationary series and that the series $\{y_t\}$ is formed from $\{x_t\}$ by the linear operator

$$y_t = \sum_{j=m}^{n} w_j x_{t-j}, \qquad \sum_{m}^{n} w_j^2 < \infty. \tag{38}$$

Such an operator is called a time-invariant linear filter. Analysis of the properties of such filters plays an important role in time series analysis since many methods of trend estimation or removal and seasonal adjustment may be represented or approximated by such filters. An interesting example that illustrates the potential pitfalls of using such filters is provided by Adelman (1965), who showed that the 20-year long swings in various economic series found by Kuznets (1961) may well have been the result of the trend filtering operations used in preliminary processing of the data. For a fuller treatment see Nerlove, Grether and Carvalho (1979, pp. 53–7).

Since the 1980s, there has been increased interest in the use of nonlinear filters for extracting the business cycle component of macroeconomic time series. Examples include the band-pass filter (Christiano and Fitzgerald, 2003) and the Hodrick–Prescott (HP) filter (Hodrick and Prescott, 1997; Ravn and Uhlig, 2002). The latter approach postulates that $y_t = \tau_t + c_t$, where τ_t denotes the trend component and c_t the deviation from trend or 'cyclical' component of the time series y_t. The trend component is chosen to minimize the loss function:

$$\sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=1}^{T} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \tag{39}$$

where $c_t = y_t - \tau_t$ and λ is a pre-specified parameter that depends on the frequency of the observations. The trade-off in this optimization problem is between the degree to which the trend component fits the data and the smoothness of the trend.

7.2. Description of seasonality and seasonal adjustment

Many economic time series exhibit fluctuations which are periodic within a year or a fraction thereof. The proper treatment of such seasonality, whether stochastic or deterministic, is the subject of a large literature, summarized rather selectively in Nerlove, Grether and Carvalho (1979, ch. 1). More recent treatments can be found in Hylleberg (1992), Franses (1996) and Ghysels and Osborn (2001).

Seasonality may be modelled and its presence detected using spectral analysis (Nerlove, 1964) or using time domain methods. Deterministic seasonality, in the form of model parameters that vary deterministically with the season, offers no great conceptual problems but many practical ones. Stochastic seasonality is often modelled in the form of seasonal unit roots. In that case, seasonal differencing of the data removes the unit root component. Multiple time series may exhibit seasonal co-integration. Sometimes it is convenient to specify stochastic seasonality in the form of an UC model (Grether and Nerlove, 1970). Appropriate UC models may be determined directly or by fitting an ARIMA model and deriving a related UC model by imposing sufficient a priori restrictions (Hillmer and Tiao, 1982; Bell and Hillmer, 1984).

7.3. Forecasting

One of the simplest forecasting procedures for time series is exponential smoothing based on the relationship

$$\hat{x}_{t+1|t} = (1 - \theta)x_t + \theta \hat{x}_{t|t-1} \tag{40}$$

where x_t, is the observed series and $\hat{x}_{j|k}$ is the forecast of the series at time j made on the basis of information available up to time k. Muth (1960) showed that (40) provides an MMSE forecast if the model generating the time series is $x_t - x_{t-1} = \varepsilon_t - \theta\varepsilon_{t-1}$. Holt (1957) and Winters (1960) generalized the exponential smoothing approach to models containing more complex trend and seasonal components. Further generalization and proofs of optimality are contained in Theil and Wage (1964) and Nerlove and Wage (1964).

Perhaps the most popular approach to forecasting time series is based on ARIMA models of time series processes (Box and Jenkins, 1970). The developments discussed in the preceding paragraph led to the development of UC models, which give rise to restricted ARIMA model forms (Nerlove, Grether and Carvalho, 1979). State-space representations of these models permit the application of the Kalman filter to both estimation and forecasting. Harvey (1984) presents a unified synthesis of the various methods.

More recently, the focus has shifted from traditional forecasting methods towards methods that exploit the increased availability of a large number of potential predictors. Consider the problem of forecasting $y_{t+h|t}$ based on its own current and past values as well as those of N additional variables, x_t. Of particular interest is the case in which the number of predictors, N, exceeds the number of time series observations, T. In that case, principal components analysis provides a convenient way of extracting a low-dimensional vector of common factors from the original data-set x_t (Stock and Watson, 2002a; 2002b). Forecasts that incorporate estimated common factors have proved successful in many cases in reducing forecast errors relative to traditional time series forecasting methods. Boivin and Ng (2005) provide a systematic comparison of alternative *factor model forecasts*. Another promising forecasting method is *Bayesian model*

averaging across alternative forecasting models (Raftery, Madigan and Hoeting, 1997). The latter method builds on the literature on forecast combinations (Bates and Granger, 1969).

7.4. Dynamic econometric modelling

There is a close connection between multivariate time-series models and the structural, reduced and final forms of dynamic econometric models; the standard simultaneous-equations model (SEM) is a specific and restricted case.

Suppose that a vector of observed variables y_t may be subdivided into two classes of variables, 'exogenous', $\{x_t\}$, and endogenous, $\{z_t\}$. A dynamic, multivariate simultaneous linear system may be written.

$$\begin{bmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ 0 & \Psi_{22}(L) \end{bmatrix} \begin{pmatrix} z_t \\ x_t \end{pmatrix} = \begin{bmatrix} \Theta_{11}(L) & 0 \\ 0 & \Theta_{22}(L) \end{bmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{41}$$

where $\Psi_{ij}(\cdot)$ and $\Theta_{ij}(\cdot), i,j = 1,2$ are matrix polynomials in the lag operator L. Such systems are known as vector ARMAX models and conditions for their identification are given by Hatanaka (1975). The reduced form of the system is obtained by expressing z_t as a function of lagged endogenous and current and lagged exogenous variables. The final form is then obtained by eliminating the lagged endogenous variables (see Zellner and Palm, 1974; Wallis, 1977).

7.5. Structural vector autoregressions

An important special case of the dynamic SEM is the structural vector autoregressive model in which all variables are presumed endogenous, the lag structure is unrestricted up to some order p, and identification of the structural from is achieved by imposing restrictions on the correlation structure of the structural innovations (Sims, 1980). The most common form of the structural VAR(p) model imposes restrictions on the contemporaneous interaction of structural innovations. Consider the structural form for a K-dimensional vector $\{x_t\}, t = 1, \ldots, T$:

$$B_0 x_t = \sum_{i=1}^{p} B_i x_{t-i} + \eta_t, \tag{42}$$

where $\eta_t \sim (0, \Sigma_\eta)$ denotes the $(K \times 1)$ vector of serially uncorrelated structural innovations (or shocks) and B_i, $i = 0, \ldots, p$, the $(K \times K)$ coefficient matrices. Without loss of generality, let $\Sigma_\eta = I$. The corresponding reduced form is

$$x_t = \sum_{i=1}^{p} B_0^{-1} B_i x_{t-i} + B_0^{-1}\eta_t = \sum_{i=1}^{p} A_i x_{t-i} + \varepsilon_t \tag{43}$$

where $\varepsilon_t \sim (0, \Sigma_\varepsilon)$. Since $\varepsilon_t = B_0^{-1}\eta_t$, it follows that $\Sigma_\varepsilon = B_0^{-1}B_0^{-1'}$. Given a consistent estimate of the reduced form parameters A_i, $i = 1, \ldots, p$, and Σ_ε, the elements of B_0^{-1}will be exactly identified after imposing $K(K-1)/2$ restrictions on the parameters of B_0^{-1} that reflect the presumed structure of the economy. Given estimates of B_0^{-1} and A_i, $i = 1, \ldots, p$, estimates of the remaining structural parameters may be recovered from $B_i = B_0 A_i$.

In practice, the number of restrictions that can be economically motivated may be smaller or larger than $K(K-1)/2$. Alternative estimation strategies that remain valid in the over-identified case include the generalized method of moments (Bernanke, 1986) and maximum likelihood (Sims, 1986). An instrumental variable interpretation of VAR estimation is discussed in Shapiro and Watson (1988). Semi-structural VAR models that are only partially identified have been proposed by Bernanke and Mihov (1998).

Alternative identification strategies may involve putting restrictions on the long-run behaviour of economic variables (Blanchard and Quah, 1989; King et al., 1991) or on the sign and/or shape of the impulse responses (Faust, 1998). Other possibilities include identification via heteroskedasticity (Rigobon, 2003) or the use of high-frequency data (Faust, Swanson and Wright, 2004).

The estimates of the structural VAR form may be used to compute the dynamic responses of the endogenous variables to a given structural shock, variance decompositions that measure the average contribution of each structural shock to the overall variability of the data, and historical decompositions of the path of x_t based on the contribution of each structural shock.

8. Conclusions

The literature on time series analysis has made considerable strides since the 1980s. The advances have been conceptual, theoretical and methodological. The increased availability of inexpensive personal computers in particular has revolutionized the implementation of time series techniques by shifting the emphasis from closed-form analytic solutions towards numerical and simulation methods. The ongoing improvements in information technology, broadly defined to include not only processing speed but also data collection and storage capabilities, are likely to transform the field even further. For example, the increased availability of large cross-sections of time series data, the introduction of ultra high-frequency data, the electronic collection of micro-level time series data (such as web-based data or scanner data), and the increased availability of data in real time all are creating new applications and spurring interest in the development of new methods of time series analysis. These developments already have brought together the fields of empirical finance and time series econometrics, resulting in the emergence of the new and fertile field of financial econometrics. As the use of time series methods

becomes more widespread in applied fields, there will be increasing interest in the development of methods that can be adapted to the specific objectives of the end user. Another question of growing importance is how to deal with rapidly evolving economic environments in the form of structural breaks and other model instabilities. Finally, the improvement of structural time series models for macroeconomic policy analysis will remain a central task if time series analysis is to retain its importance for economic policymaking.

FRANCIS X. DIEBOLD, LUTZ KILIAN AND MARC NERLOVE

See also **cointegration; factor models; forecasting; long memory models; maximum likelihood; nonlinear time series analysis; seasonal adjustment; spectral analysis; spurious regressions; trend/cycle decomposition; vector autoregressions.**

Bibliography

Adelman, I. 1965. Long cycles – fact or artifact? *American Economic Review* 60, 443–63.

Akaike, H. 1970. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics* 22, 203–17.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–23.

Andersen, T., Bollerslev, T., Christoffersen, P. and Diebold, F. 2006. Volatility and correlation forecasting. In *Handbook of Economic Forecasting*, ed. G. Elliott, C. Granger and A. Zimmermann. Amsterdam: North-Holland.

Andersen, T., Bollerslev, T. and Diebold, F. 2006. Parametric and nonparametric volatility measurement. In *Handbook of Financial Economics*, ed. L. Hansen and Y. Ait-Sahalia. Amsterdam: North-Holland.

Anderson, T. 1971. *The Statistical Analysis of Time Series.* New York: John Wiley.

Anderson, T. 1977. Estimation for autoregressive moving average models in the time and frequency domains. *Annals of Statistics* 5, 842–65.

Anderson, T. and Takemura, A. 1986. Why do noninvertible moving averages occur? *Journal of Time Series Analysis* 7, 235–54.

Ansley, C. 1979. An algorithm for the exact likelihood of a mixed autoregressive-moving average process. *Biometrika* 66, 59–65.

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

Baillie, R., Bollerslev, T. and Mikkelsen, H.-O. 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3–30.

Bates, J. and Granger, C. 1969. The combination of forecasts. *Operational Research Quarterly* 20, 451–68.

Bell, W. and Hillmer, S. 1984. Issues involved with seasonal analysis of economic time series. *Journal of Business and Economic Statistics* 2, 291–349.

Bernanke, B. 1986. Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy* 25, 49–100.

Bernanke, B. and Mihov, I. 1998. Measuring monetary policy. *Quarterly Journal of Economics* 113, 869–902.

Beveridge, W. 1921. Weather and harvest cycles. *Economic Journal* 31, 429–52.

Beveridge, W. 1922. Wheat prices and rainfall in western Europe. *Journal of the Royal Statistical Society* 85, 412–59.

Blanchard, O. and Quah, D. 1989. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review* 79, 655–73.

Boivin, J. and Ng, S. 2005. Understanding and comparing factor-based forecasts. *International Journal of Central Banking* 1(3), 117–52.

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

Box, G., Hillmer, S. and Tiao, G. 1978. Analysis and modeling of seasonal time series. In *Seasonal Analysis of Economic Time Series.* ed. A. Zellner. Washington, DC: Bureau of the Census, Department of Commerce.

Box, G. and Jenkins, G. 1970. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Brillinger, D. 1975. *Time Series: Data Analysis and Theory.* New York: Holt.

Burman, J. 1980. Seasonal adjustment by signal extraction. *Journal of the Royal Statistical Society. Series A* 143, 321–37.

Burridge, P. and Wallis, K. 1985. Calculating the variance of seasonally adjusted series. *Journal of the American Statistical Association* 80, 541–52.

Cappé, O., Moulines, E. and Ryden, T. 2005. *Inference in Hidden Markov Models.* New York: Springer-Verlag.

Chambers, M. 1998. Long memory and aggregation in macroeconomic time series. *International Economic Review* 39, 1053–72.

Christiano, L. and Fitzgerald, T. 2003. The band pass filter. *International Economic Review* 44, 435–65.

Christoffersen, P. and Diebold, F. 1996. Further results on forecasting and model selection under asymmetric loss. *Journal of Applied Econometrics* 11, 561–71.

Christoffersen, P. and Diebold, F. 1997. Optimal Prediction under asymmetric loss. *Econometric Theory* 13, 808–17.

Cournot, A. 1838. Researches into the Mathematical Principles of the Theory of Wealth, trans. N. Bacon. New York: Macmillan, 1927.

Diebold, F. and Inoue, A. 2001. Long memory and regime switching. *Journal of Econometrics* 105, 131–59.

Doob, J. 1953. *Stochastic Processes.* New York: John Wiley.

Durbin, J. and Koopman, S. 2001. *Time Series Analysis by State Space Methods.* Oxford: Oxford University Press.

Durlauf, S. and Phillips, P. 1988. Trends versus random walks in time series analysis. *Econometrica* 56, 1333–54.

Elliott, G. 1998. The robustness of co-integration methods when regressors almost have unit roots. *Econometrica* 66, 49–58.

Engle, R. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica* 50, 987–1008.

Engle, R. and Granger, C. 1987. Co-integration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–76.

Engle, R. and Granger, C. 1991. *Long Run Economic Relations: Readings in Co-integration.* Oxford: Oxford University Press.

Faust, J. 1998. The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy* 49, 207–44.

Faust, J., Swanson, E. and Wright, J. 2004. Identifying VARs based on high frequency futures data. *Journal of Monetary Economics* 51, 1107–31.

Fishman, G. 1969. *Spectral Methods in Econometrics.* Cambridge: Harvard University Press.

Franses, P. 1996. *Periodicity and Stochastic Trends in Economic Time Series.* Oxford: Oxford University Press.

Fuller, W. 1976. *Introduction to Statistical Time Series.* New York: John Wiley.

Geweke, J. and Porter-Hudak, S. 1983. The estimation and application of long memory series models. *Journal of Time Series Analysis* 4, 221–38.

Ghysels, E. and Osborn, D. 2001. *The Econometric Analysis of Seasonal Time Series.* Cambridge: Cambridge University Press.

Granger, C. 1966. The typical spectral shape of an economic variable. *Econometrica* 34, 150–61.

Granger, C. 1969. Prediction with a generalized cost of error function. *Operations Research Quarterly* 20, 199–207.

Granger, C. 1980. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14, 227–38.

Granger, C. 1981. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* 16, 121–30.

Granger, C. and Newbold, P. 1974. Spurious regressions in econometrics. *Journal of Econometrics* 2, 111–20.

Granger, C. and Newbold, P. 1977. *Forecasting Economic Time Series.* New York: Academic Press.

Granger, C. and Teräsvirta, T. 1993. *Modelling Nonlinear Economic Relationships.* Oxford: Oxford University Press.

Grether, D. and Nerlove, M. 1970. Some properties of 'optimal' seasonal adjustment. *Econometrica* 38, 682–703.

Hamilton, J. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357–84.

Hannan, E. 1960. *Time Series Analysis.* London: Methuen.

Hannan, E. 1969. The estimation of mixed moving average autoregressive systems. *Biometrika* 56, 223–5.

Hannan, E. 1970. *Multiple Time Series.* New York: John Wiley.

Hannan, E. 1971. The identification problem for multiple equation systems with moving average errors. *Econometrica* 39, 751–65.

Hannan, E. 1976. The identification and parameterization of ARMAX and state space forms. *Econometrica* 44, 713–23.

Hannan, E. 1979. The statistical theory of linear systems. In *Developments in Statistics*, ed. P. Krishnaiah. New York: Academic Press.

Harvey, A. 1981. *Time Series Models.* Oxford: Allan.

Harvey, A. 1984. A unified view of statistical forecasting procedures. *Journal of Forecasting* 3, 245–75.

Harvey, A. 1989. *Forecasting. Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Harvey, A. and Peters, S. 1990. Estimation procedures for structural time series models. *Journal of Forecasting* 9, 89–108.

Harvey, A. and Todd, P. 1984. Forecasting economic time series with structural and Box–Jenkins models: a case study (with discussion). *Journal of Business and Economic Statistics* 1, 299–315.

Hatanaka, M. 1975. On the global identification of the dynamic simultaneous equations model with stationary disturbances. *International Economic Review* 16, 545–54.

Hillmer, S., Bell, W. and Tiao, G. 1983. Modeling considerations in the seasonal analysis of economic time series. In *Applied Time Series Analysis of Economic Data*, ed. A. Zellner. Washington, DC: Bureau of the Census, Department of Commerce.

Hillmer, S. and Tiao, G. 1982. An ARIMA-model-based approach to seasonal adjustment. *Journal of the American Statistical Association* 77, 63–70.

Hodrick, R. and Prescott, E. 1997. Postwar US business cycles: an empirical investigation. *Journal of Money, Credit and Banking* 29, 1–16.

Holt, C. 1957. Forecasting seasonals and trends by exponentially weighted moving averages. ONR Research Memorandum No. 52, Carnegie Institute of Technology.

Hylleberg, S. 1992. *Modelling Seasonality.* Oxford: Oxford University Press.

Jevons, W. 1884. *Investigations in Currency and Finance.* London: Macmillan.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford: Oxford University Press.

Kalman, R. 1960. A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers – Journal of Basic Engineering* D 82, 35–45.

King, R., Plosser, C., Stock, J. and Watson, M. 1991. Stochastic trends and economic fluctuations. *American Economic Review* 81, 819–40.

Kolmogorov, A. 1941. Interpolation und Extrapolation von stationären zufälligen Folgen. *Bulletin of the Academy Science (Nauk), USSR, Mathematical Series* 5, 3–14.

Koopmans, L. 1974. *The Spectral Analysis of Time Series.* New York: Academic Press.

Kuznets, S. 1961. *Capital and the American Economy: Its Formation and Financing*. New York: Princeton University Press for the National Bureau of Economic Research.

Maravall, A. 1981. Desestacionalization y Politica Monetaria. Economic Studies No. 19. Madrid: Bank of Spain.

Maravall, A. 1984. Model-based treatment of a manic depressive series. Working paper, Bank of Spain.

Muth, J. 1960. Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association* 55, 299–305.

Nelson, C. and Kang, H. 1981. Spurious periodicity in inappropriately detrended time series. *Econometrica* 49, 741–52.

Nelson, C. and Plosser, C. 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics* 10, 139–62.

Nerlove, M. 1964. Spectral analysis of seasonal adjustment procedures. *Econometrica* 32, 241–86.

Nerlove, M. 1965. A comparison of a modified Hannan and the BLS seasonal adjustment filters. *Journal of the American Statistical Association* 60, 442–91.

Nerlove, M. 1967. Distributed lags and unobserved components in economic time series. In *Ten Economic Essays in the Tradition of Irving Fisher*, ed. W. Fellner et al. New York: John Wiley.

Nerlove, M., Grether, D. and Carvalho, J. 1979. *Analysis of Economic Time Series*. New York: Academic Press.

Nerlove, M. and Wage, S. 1964. On the optimality of adaptive forecasting. *Management Science* 10, 207–24.

Newbold, P. 1974. The exact likelihood function for a mixed autoregressive-moving average process. *Biometrika* 61, 423–6.

Parzen, E. 1961. An approach to time series analysis. *Annals of Mathematical Statistics* 32, 951–89.

Phillips, P. 1986. Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–40.

Phillips, P. and Durlauf, S. 1986. Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473–95.

Phillips, P. and Hansen, B. 1990. Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Pierce, D. 1978. Seasonal adjustment when both deterministic and stochastic seasonality are present. In *Seasonal Analysis of Economic Time Series*, ed. A. Zellner. Washington, DC: Bureau of the Census, Department of Commerce.

Pierce, D. 1979. Signal extraction error in nonstationary time series. *Annals of Statistics* 7, 1303–20.

Priestley, M. 1981. *Spectral Analysis and Time Series*. New York: Academic Press.

Raftery, A., Madigan, D. and Hoeting, J. 1997. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association* 92, 179–91.

Ravn, M. and Uhlig, H. 2002. On adjusting the HP-filter for the frequency of observations. *Review of Economics and Statistics* 84, 371–76.

Rigobon, R. 2003. Identification through heteroskedasticity. *Review of Economics and Statistics* 85, 777–92.

Rudebusch, G. 1993. The uncertain unit root in real GNP. *American Economic Review* 83, 264–72.

Sargan, J. and Bhargava, A. 1983. Maximum likelihood estimation of regression models with moving average errors when the root lies on the unit circle. *Econometrica* 51, 799–820.

Schuster, A. 1898. On the investigation of hidden periodicities with application to the supposed 26-day period of meteorological phenomena. *Terrestrial Magnetism and Atmospheric Electricity* [now *Journal of Geophysical Research*] 3, 13–41.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6, 461–4.

Schweppe, F. 1965. Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory* 11, 61–70.

Shapiro, M. and Watson, M. 1988. Sources of business cycle fluctuations. *NBER Macroeconomics Annual* 3, 111–56.

Shephard, N. 2005. *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

Sims, C. 1980. Macroeconomics and reality. *Econometrica* 48, 1–48.

Sims, C. 1986. Are forecasting models usable for policy analysis? *Quarterly Review, Federal Reserve Bank of Minneapolis* 10, 2–16.

Sims, C., Stock, J. and Watson, M. 1991. Inference in linear time series models with some unit roots. *Econometrica* 58, 113–44.

Slutsky, E. 1927. The summation of random causes as the source of cyclic processes. *Econometrica* 5, April 1937, 105–46.

Sowell, F. 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53, 165–88.

Stock, J. 1991. Confidence intervals for the largest autoregressive root in US economic time series. *Journal of Monetary Economics* 28, 435–60.

Stock, J. and Watson, M. 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.

Stock, J. and Watson, M. 2002a. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–79.

Stock, J. and Watson, M. 2002b. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147–62.

Stokes, G. 1879. Note on searching for hidden periodicities. *Proceedings of the Royal Society* 29, 122–5.

Theil, H. and Wage, S. 1964. Some observations on adaptive forecasting. *Management Science* 10, 198–206.

Wallis, K. 1977. Multiple time series analysis and the final form of econometric models. *Econometrica* 45, 1481–97.

Whittle, P. 1963. *Prediction and Regulation by Linear Least-squares Methods*. London: English Universities Press.
Wiener, N. 1949. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. New York: John Wiley.
Winters, P. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science* 6, 324–42.
Wold, H. 1938. *A Study in the Analysis of Stationary Time Series*. Stockholm: Almqvist and Wiksell.
Yule, G. 1921. On the time-correlation problem, with special reference to the variate-difference correlation method. *Journal of the Royal Statistical Society* 84, 497–526.
Yule, G. 1926. Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series. *Journal of the Royal Statistical Society* 89, 1–64.
Yule, G. 1927. On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London A* 226, 267–98.
Zellner, A. and Palm, F. 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* 2, 17–54.

time use

Time is a finite, irreplaceable resource available to every man, woman and child in equal amounts of 168 hours per week over the course of life. Time use refers to the allocation of time to alternative uses such as sleep, leisure or work.

Time is perhaps the most fundamental scarce resource; unlike money income or wealth, it is equally distributed; and how well or wastefully it is used largely determines the progress, achievement and well-being of individuals, families, communities and societies.

One way to assess progress and well-being is to measure aggregate changes in the uses of time rather than, or in addition to, the usual monetary statistics of national income and expenditure. National time accounts are more comprehensive than money accounts as they simultaneously measure the productive time spent in both the market and the household, as well as the time spent in consumption of outputs from both.

The advent of radio astronomy in 1957, using a wider spectrum than that provided by the visible light frequencies, opened up new views of the astronomical universe. Similarly, official statistical organizations are beginning to provide new views of the economic and social universe by observing the world through a 'time frequency' rather than the more visible 'money frequency'.

Researching time use and human behaviour

Time use economics is the scientific study of human behaviour concerning the allocation of time between alternative uses (cf. Robbins, 1932, p. 16).

Time use can be usefully considered under the three comprehensive categories of production, consumption and investment. These accord with the theoretical framework of macroeconomics. Time use data show how individuals and households allocate their time between production (paid work and unpaid work), consumption (meals, television, social interaction and recreation) and human capital formation and maintenance (education, self-care and sleep).

Margaret Reid provided the well-known 'third person' test for separating production from consumption: 'If an activity is of such character that it might be delegated to a paid worker, then that activity shall be deemed productive' (Reid, 1934, p. 11).

To facilitate analysis, macroeconomics makes a further distinction between resources for immediate use – consumption – and those for use over a longer time period – investment. Accordingly, time spent learning a skill or gaining knowledge in education is clearly investment in human capital. Similarly, time spent in sleep and self-care can be regarded as a necessary daily investment to maintain functioning minds and bodies.

When people spend time, in economic terms they are effectively allocating time between market production, household production, consumption and investment. Modelling of household time-allocation decisions goes beyond understanding the simple work-leisure trade-off. It provides knowledge of the detailed interactions between production, consumption and investment activities. It also provides a framework for the analysis of the derived demands for market commodities implicit in household production and consumption.

In a much-quoted article in *Scientific American*, Vanek (1974) surveys the time spent in household production in the United States over a period of 40 years. Juster and Stafford (1991) provides a comprehensive appraisal of the importance of time allocation as an analytic construct and a review of what had been learned from time allocation data in modelling economic behaviour and the dynamics of economic change. Robinson and Godbey (1999) give an account of the American data from 1965, 1975, 1985 and 1995, and Gershuny (2000) comprehensively surveys the time use data for Britain.

Becker (1965) offers a theoretical framework for the microeconomics of time use by including the cost of time on the same footing as the cost of market goods and by assuming households 'combine time and market goods to produce more basic commodities that directly enter their utility functions' (1965, p. 494). This neoclassical approach to a theory has been criticized on many grounds, summarized recently by Folbre (2004).

While neoclassical micro-theory provides a useful framework for a *microeconomics of household production and consumption*, it completely disregards the full macroeconomic magnitude of household production. In most developed countries more labour is required for household production than for market production