the 20th century, such proposals have had little practical impact. A unit of account is an economy-wide institution, and it appears to be both difficult and costly to bring about a collective decision to replace even a badly functioning one with a specific alternative. Nevertheless, in conditions of very rapid inflation, money's means of exchange and unit of account functions do tend to become separated, albeit more often in a piecemeal fashion than as the outcome of any coherent policy decision. As inflation rises, some stable foreign currency comes to be more and more widely used as a unit of account, even when actual transactions are still completed with local currency valued at the current exchange rate. The recent widespread use of the US dollar as a unit of account in Israel and several Latin American countries has given rise to the word 'dollarization' to describe this phenomenon.

DAVID LAIDLER

See also CURRENCY SUBSTITUTION; EXOTIC CURRENCIES; LEGAL TENDER; MENU COSTS; MONETARY CONSTITUTIONS; MONEY; NEW MONETARY ECONOMICS.

BIBLIOGRAPHY

Cowen, T. and Kroszner, R. 1987. The development of the new monetary economics. *Journal of Political Economy* 95: 567–90.

Einaudi, L. 1953. The theory of imaginary money from Charlemagne to the French Revolution. In *Enterprise and Secular Change*, ed. F. Lane and L. Riemersma, Homewood, Ill.: Richard Irwin.

Fisher, I. 1911. *The Purchasing Power of Money*. New York: Macmillan.

Jevons, W.S. 1875. *Money and the Mechanism of Exchange*. London: King.

Marshall, A. 1887. Remedies for fluctuations in general prices. *Contemporary Review*. In *Memorials of Alfred Marshall*, ed. A.C. Pigou, London: Macmillan 1925.

Rotemberg, J.J. 1987. The new-Keynesian microfoundations. *NBER Macroeconomics Annual* 2: 69–104.

unit roots. Economic and financial time series are frequently well modelled by autoregressive moving-average (ARMA) schemes of the type

$$a(L)y_t = b(L)\varepsilon_t;$$

$$a(L) = \sum_0^p a_i L^i,\ b(L) = \sum_0^q b_j L^j, \tag{1}$$

where ε_t is an orthogonal sequence (i.e. $E(\varepsilon_t) = 0$, $E(\varepsilon_t\varepsilon_s) = 0$ for all $t \neq s$), L is the backshift operator for which $Ly_t = y_{t-1}$ and $a(L)$, $b(L)$ are finite-order lag polynomials whose leading coefficients are $a_0 = b_0 = 1$. Parsimonious schemes (often with $p+q \leq 3$) are usually selected in practice either by informal 'model identification' processes such as those described in the text by Box and Jenkins (1976) or more formal order-selection criteria which penalize choices of large p and/or q. Model (1) is assumed to be irreducible, so that $a(L)$ and $b(L)$ have no common factors. The model (1) and time series y_t are said to have an *autoregressive unit root* if $a(L)$ factors as $(1-L)a_1(L)$ and a *moving-average unit root* if $b(L)$ factors as $(1-L)b_1(L)$.

Much attention has been focused recently on models with autoregressive unit roots. If the polynomial $a_1(L)$ has all its zeros outside the unit circle then we may write (1) as

$$\Delta y_t = u_t = a_1(L)^{-1}b(L)\varepsilon_t. \tag{2}$$

This suggests more general models where, for instance, u_t may be a linear process of the form

$$u_t = c(L)\varepsilon_t = \sum_0^\infty c_j\varepsilon_{t-j},\ \text{with} \sum_0^\infty c_j^2 < \infty, \tag{3}$$

or a general stationary process with spectrum $f_u(\lambda)$. If we solve (2) with an initial state at $t = 0$ we have the important representation

$$y_t = \sum_1^t u_j + y_0 = S_t + y_0, \tag{4}$$

showing that S_t and hence y_t are 'accumulated' or 'integrated' processes. A time series y_t that satisfies (2) is therefore said to be integrated of order one (i.e. it has one unit root or is an I(1) process) provided $f_u(0) > 0$. The latter condition rules out the possibility of a moving-average unit root in the model for u_t that would cancel the effect of the autoregressive unit root (e.g. if $b(L) = (1-L)b_1(L)$ then model (2) is $\Delta y_t = \Delta a_1(L)^{-1}b_1(L)\varepsilon_t$ or, after cancellation, just $y_t = a_1(L)^{-1}b_1(L)\varepsilon_t$ which is not I(1)). Note that this possibility is also explicitly ruled out in the ARMA case by the requirement that $a(L)$ and $b(L)$ have no common factors. The process S_t in (4) is often described as a *stochastic trend*.

The representation (4) is especially important because it shows that the effect of the random shocks u_j on y_t do not die out as $t-j$ grows large. This means that the shocks u_j have a *persistent* effect on y_t in this model, in contrast to stationary systems. Whether actual economic time series have this characteristic or not is of course an empirical issue. The question can be addressed through statistical tests for the presence of a unit root in the series, a subject which has grown to be of major importance in recent years and which will be discussed later in this essay. From the perspective of economic modelling the issue of persistence is also important because if macroeconomic variables like real GNP have a unit root then shocks to real GNP have permanent effects, whereas in traditional business cycle theory the effect of shocks on real GNP is usually considered to be only temporary. In more recent real business cycle theory variables like real GNP are modelled in such a way that in the long run their paths are determined by supply side shocks that can be ascribed to technological and demographic forces from outside the model. Such economic models are more compatible with the statistical model (4).

Permanent and transitory effects in (4) can be distinguished by decomposing the process u_t in (3) as follows:

$$\begin{aligned} u_t &= \{C(1) + (L-1)\tilde{C}(L)\}\varepsilon_t \\ &= C(1)\varepsilon_t + \tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t, \end{aligned} \tag{5}$$

where $\tilde{\varepsilon}_t = \tilde{C}(L)\varepsilon_t$, $\tilde{C}(L) = \Sigma_0^\infty \tilde{c}_j L^j$ and $\tilde{c}_j = \Sigma_{j+1}^\infty c_s$. The decomposition (5) is valid algebraically if $\Sigma_0^\infty j^{1/2}|c_j| < \infty$.

Equation (5) is called the Beveridge–Nelson (1981) or BN decomposition of u_t. When it is applied to (4) it yields the representation

$$y_t = C(1)\sum_1^t \varepsilon_j + \tilde{\varepsilon}_0 - \tilde{\varepsilon}_t + y_0$$

$$= C(1)\sum_1^t \varepsilon_j + \xi_t. \qquad (6)$$

Under (6) we have a direct decomposition of y_t into two components, one where the effects of shocks are permanent, viz. $C(1)\Sigma_1^t\varepsilon_j$, and the other where the effects of shocks are transitory, viz. $\xi_t = \tilde{\varepsilon}_0 - \tilde{\varepsilon}_t + y_0$ (since $\tilde{\varepsilon}_t$ is stationary). The first component is the martingale or random walk component of y_t and its relative strength is measured by the magnitude of the coefficient $C(1)$.

Model (4) is of special interest to economists working in finance because its output, y_t, behaves as if it has no fixed mean and this is a characteristic of many financial time series. Indeed, if the components u_j are independent and identically distributed (iid) then y_t is a random walk. More generally, if u_j is a martingale difference sequence (mds) (i.e. orthogonal to its own past history so that $E_{j-1}(u_j) = E(u_j|u_{j-1}, u_{j-2}, \ldots, u_1) = 0$) then y_t is a martingale. Martingales are the essential mathematical elements in the development of a theory of fair games and they now play a key role in the mathematical theory of finance, exchange rate determination and securities markets. Duffie (1988) provides a modern treatment of finance that makes extensive use of this theory.

In empirical finance much attention has recently been given to models where the conditional variance $E(u_j^2|u_{j-1}, u_{j-2}, \ldots, u_1) = \sigma_j^2$ is permitted to be time varying. Such models have been found to fit financial data well and many different parametric schemes for σ_j^2 have been devised, of which the ARCH (autoregressive conditional heteroskedasticity) and GARCH (generalized ARCH) models are the most common. These models come within the general class of models like (1) with mds errors. Some models of this kind also allow for the possibility of a unit root in the determining mechanism of the conditional variance σ_j^2 and these are called integrated conditional heteroskedasticity models. The IGARCH (integrated GARCH) model of Engle and Bollerslev (1986) is an example, where for $\omega \geqslant 0, \beta \geqslant 0$ and $\alpha > 0$ we have

$$\sigma_j^2 = \omega + \beta\sigma_{j-1}^2 + \alpha u_{j-1}^2;$$
$$u_j = \sigma_j z_j,\ \alpha + \beta = 1;$$
$$\sigma_j^2 = \omega + \sigma_{j-1}^2 + \alpha\sigma_{j-1}^2(z_{j-1}^2 - 1), \qquad (7)$$

and the innovations z_j are iid with $E(z_j) = 0$ and $E(z_j^2) = 1$. It is apparent from (7) that this model for σ_j^2 has an autoregressive unit root and, indeed, since

$$\mathrm{E}(\sigma_j^2|\sigma_{j-1}^2) = \omega + \sigma_{j-1}^2,$$

σ_j^2 is a martingale when $\omega = 0$. It is also apparent from (7) that shocks as manifested in the deviation $z_{j-1}^2 - 1$ are persistent in σ_j^2. Thus σ_j^2 shares some of the characteristics of an I(1) integrated process. But in other ways, σ_j^2 is very different. For instance, when $\omega = 0$ then $\sigma_j^2 \to 0$ almost surely and when $\omega > 0$, σ_j^2 is asymptotically equivalent to a strictly stationary and ergodic process. These and other features of models like (4) for conditional variance processes with a unit root are studied in Nelson (1990).

In macroeconomic theory also, models such as (2) play a central role in modern treatments. In a highly influential paper, Hall (1978) showed that under some general conditions consumption is well modelled as a martingale, so that consumption in the current period is the best predictor of future consumption, thereby providing a macroeconomic version of the efficient markets hypothesis. Much attention has been given to this idea in subsequent empirical work.

One generic class of economic model where unit roots play a special role is the 'present value model' of Campbell and Shiller (1988). This model is based on agents' forecasting behaviour and takes the form of a relationship between one variable Y_t and the discounted, present value of rational expectations of future realizations of another variable $X_{t+i} (i = 0, 1, 2, \ldots)$. More specifically, for some stationary sequence c_t (possibly a constant) we have

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty} \delta^i E_t(X_{t+i}) + c_t. \qquad (8)$$

When X_t is a martingale, $E_t(X_{t+i}) = X_t$ and (8) becomes

$$Y_t = \theta X_t + c_t \qquad (9)$$

so that Y_t and X_t are cointegrated in the sense of Engle and Granger (1987). More generally, when X_t is I(1) we have

$$Y_t = \theta X_t + \bar{c}_t,$$

where $\bar{c}_t = c_t + \theta\Sigma_{k=1}^{\infty}\delta^k E_t(\Delta X_{t+k})$, so that Y_t and X_t are also cointegrated in this general case. Models of this type arise naturally in the study of the term structure of interest rates, stock prices and dividends and linear-quadratic intertemporal optimization problems.

Statistical tests for the presence of a unit root fall into the two general categories of classical and Bayesian, corresponding to the mode of inference that is employed. Most empirical work to date has used classical methods but attention has very recently shifted to Bayesian alternatives. Both approaches will be discussed in what follows.

Classical tests for a unit root may be classified into parametric, semiparametric and nonparametric categories. Parametric tests usually rely on augmented regressions of the type

$$\Delta y_t = a y_{t-1} + \sum_1^{k-1} \varphi_i \Delta y_{t-i} + \varepsilon_t, \qquad (10)$$

where the lagged variables are included to model the stationary error u_t in (2). Under the null hypothesis of a unit root, we have $a = 0$ in (10) whereas when y_t is stationary we have $a < 0$. Thus, a simple test for the presence of a unit root against a stationary alternative in (2) is based on a one-sided t-ratio test of H_0: $a = 0$ against H_1: $a < 0$. This test is popularly known as the ADF (or augmented Dickey–Fuller) test. It has been extensively used in empirical econometric work since the Nelson and Plosser (1982) study, where it was applied to 14 historical time series for the USA leading to the conclusion that unit roots could not be rejected for 13 of these series (all but the unemployment rate). In that study the alternative hypothesis was that the series were stationary about a deterministic trend (i.e. trend stationary) and therefore model (10) was further augmented to include a linear trend, viz.

$$\Delta y_t = \mu + \beta t + a y_{t-1} + \sum_1^{k-1} \varphi_i \Delta y_{t-i} + \varepsilon_t. \qquad (10')$$

When y_t is trend stationary we have $a<0$ and $\beta \neq 0$ in (10′), so the null hypothesis of a difference stationary process is $a = 0$ and $\beta = 0$. This null hypothesis allows for the presence of a non-zero drift in the process when the parameter $\mu \neq 0$. In this case a joint test of the null hypothesis H_0: $a = 0$, $\beta = 0$ can be mounted using a regression F-test. What distinguishes both this test and the corresponding t-test in (10) is that critical values for these tests are not the same as those for conventional regression F- and t-tests, even in large samples. Under the null, the limit theory for these tests involves functionals of a Wiener process and typically the critical values for five or one percent level tests are much further out than those of the standard normal or chi-squared distributions. The limit theory was first explored and tabulations provided by Dickey (1976), Fuller (1976) and Dickey and Fuller (1979, 1981). Later work by Said and Dickey (1984) showed that if the lag number k in (10) is allowed to increase as the sample size increases then the ADF test is asymptotically valid.

Several other parametric procedures have been suggested including instrumental variable methods (Hall 1989; Phillips and Hansen 1990) and variable addition methods (Park 1990). The latter also allow a null hypothesis of trend stationarity to be tested directly, rather than as an alternative to difference stationarity. Another approach that provides a test of a null of trend stationarity is based on the unobserved components representation

$$y_t = \mu + \beta t + r_t + u_t,\ r_t = r_{t-1} + v_t, \qquad (11)$$

which decomposes a time series y_t into a deterministic trend, an integrated process or random walk (r_t) and a stationary residual (u_t). The presence of the integrated process component in y_t can then be tested by testing whether the variance (σ_v^2) of the innovation v_t is zero. The null hypothesis is then H_0: $\sigma_v^2 = 0$ and this hypothesis can be simply tested using the Lagrange multiplier (LM) principle as in Kwiatkowski, Phillips and Schmidt (1990).

By combining r_t and u_t in (11) the components model may also be written as

$$y_t = \mu + \beta t + x_t, \quad \Delta x_t = a x_{t-1} + \eta_t. \qquad (11')$$

In this format it is easy to construct an LM test of the null hypothesis that y_t has a stochastic trend component by testing whether $a = 0$ in (11′). When $a = 0$, (11′) reduces to

$$\Delta y_t = \beta + \eta_t \quad \text{or} \quad y_t = \beta t + \sum_1^t \eta_j + y_0, \qquad (11'')$$

and so the parameter μ is irrelevant (or surplus) under the null. However, the parameter β retains the same meaning as the deterministic trend term coefficient under both the null and the alternative hypothesis. This approach has formed the basis of several tests for a unit root that have been developed (see Bhargava 1986 and Schmidt and Phillips 1989) and the parameter economy of this model gives these tests some advantage in terms of power over procedures like the ADF in the neighbourhood of the null.

Semiparametric unit root tests use nonparametric methods to model and to estimate the contribution from the stationary error u_t in (2). Direct least squares regression on the equation

$$\Delta y_t = a y_{t-1} + u_t \qquad (10'')$$

gives an estimate of the coefficient and its t-ratio in this equation. These statistics are then corrected to deal with serial correlation in u_t by employing an estimate of the variance of u_t and its long-run variance, which is the value of the spectrum of u_t at the zero frequency. The latter estimate may be obtained by a variety of kernel-type spectral estimates using the residuals $\hat{u}_t$ of the OLS regression on (10″). This semiparametric approach was introduced in Phillips (1987) and leads to two test statistics from (10″), one based on the coefficient estimate and called the $Z(a)$ test, the other based on its t-ratio and called the $Z(t)$ test. the $Z(t)$ test and ADF test are asymptotically equivalent. Semiparametric corrections can also be applied to the components models (11) and (11′) leading to generally applicable LM tests of stationarity ($\sigma_v^2 = 0$) and stochastic trends ($a = 0$).

More general nonparametric tests for a unit root are also possible. These rely on frequency domain regressions on (10″) over all frequency bands. They may be regarded as fully nonparametric because they test in a general way for coherency between the series y_t and its first difference Δy_t.

The $Z(a)$, $Z(t)$ and ADF tests are the most commonly used tests in empirical research. Extensive simulations have been conducted to evaluate the performance of the tests. It is known that the $Z(a)$, $Z(t)$ and ADF tests all perform satisfactorily except when the error process u_t displays strong negative serial correlation. The $Z(a)$ test generally has greater power than the other two tests but also suffers from more serious size distortion. All of these tests can be used to test for the presence of cointegration by using the residuals from a cointegrating regression. Modification of the critical values used in these tests is then required – see Phillips and Ouliaris (1990).

Attention has also been given to models like (10′) where there are structural breaks in the intercept (μ) and trend coefficient (β). Such models attach unit weight and hence persistence to the effects of innovations at particular times in the sample period. To the extent that persistent shocks of this type occur intermittently throughout the entire history of a process, these models are very similar to models with a stochastic trend. However, if only one such break occurs (or more generally a finite number of breaks) then the process does have different characteristics from that of a stochastic trend. Attempts to distinguish empirically between models with stochastic trends and models with structural breaks have been performed but with mixed results that depend on assumptions made about the timing of the break points. If the break points are known in advance, as in Perron's (1989) treatment (viz. the 1929 crash and the 1974 oil shock), both theory and application are simple, but the empirical results obtained are often not robust to the more important case where the break points are permitted to be random (e.g. Christiano 1988, Zivot and Andrews 1990).

Most empirical work on unit roots has relied on classical tests of the type described above. But Bayesian methods are also available and appear to offer certain advantages like an exact finite sample analysis. In addressing the problem of trend determination, traditional Bayes methods may be employed such as the computation of Bayesian confidence sets and the use of posterior odds tests. In both cases prior distributions on the parameters of the model need to be defined and posteriors can be calculated either by analytical methods or by numerical integration. If (10) is rewritten as

$$y_t = \rho y_{t-1} + \sum_1^{k-1} \varphi_i \Delta y_{t-i} + \varepsilon_t \qquad (10''')$$

then the posterior probability of the nonstationary set $\{\rho \geqslant 1\}$ is of special interest in assessing the evidence in support of the presence of a stochastic trend in the data. Posterior odds tests typically proceed with 'spike and slab' prior distributions (π) that assign an atom of mass such as $\pi(\rho = 1) = \theta$ to the unit-root null and a continuous distribution with mass $1 - \theta$ to the stationary alternative, so that $\pi(\rho < 1) = 1 - \theta$. The posterior odds then show how the prior odds ratio $\theta/(1-\theta)$ in favour of the unit root is updated by the data. Clearly, the input of information via the prior distribution, whether deliberate or unwitting, is a major reason for potential divergence between Bayesian and classical statistical analyses. Methods of setting an objective correlative in Bayesian analysis through the use of model-based, impartial reference priors that accommodate nonstationarity are therefore of substantial interest. These are explored in Phillips (1991a).

Empirical illustrations of the use of Bayesian methods of trend determination for various macroeconomic and financial time series are given in DeJong and Whiteman (1991a, 1991b), Schotman and Van Dijk (1991) and Phillips (1991a), the latter implementing an objective model-based approach. Most recently, Phillips and Ploberger (1991) have developed a class of 'Bayes model' tests (including a posterior odds test) that take account of the fact that Bayesian time series analysis is conducted conditionally on the realized history of the process. The mathematical effect of such conditioning is to translate models such as (10''') to a 'Bayes model' with time-varying and data-dependent coefficients, that is,

$$y_{t+1} = \hat{\rho}_t y_t + \sum_1^{k-1} \hat{\varphi}_{it} \Delta y_{t-i} + \varepsilon_{t+1}, \qquad (10'''')$$

where $(\hat{\rho}_t, \hat{\varphi}_{it};\ i = 1, \ldots, k-1)$ are the latest best estimates of the coefficients from the data available to point 't' in the trajectory. The 'Bayes model' (10'''') and its probability measure can be used to construct likelihood ratio and posterior odds tests of hypotheses such as the unit root null $\rho = 1$. Some empirical illustrations of this approach are given in Phillips (1991b, 1991c).

Nonstationarity is certainly one of the most dominant and enduring characteristics of macroeconomic and financial time series. It therefore seems appropriate that this feature of the data be seriously addressed both in econometric methodology and in empirical practice. However, until recently this has not been the case. Before 1980, it was standard empirical practice in econometrics to treat observed trends as simple deterministic functions of time. Nelson and Plosser (1982) challenged this practice and showed that observed trends are better modelled if one allows for stochastic trends. Since their work there has been a continuing reappraisal of trend behaviour in economic time series and substantial development in the econometric methods of nonstationary time series. This essay has touched only a part of this large research field and traced only the main ideas involved in unit root modelling and statistical testing. The reader may consult recent reviews by Diebold and Nerlove (1990), Dolado *et al.* (1990) and Campbell and Perron (1991) and special issues of the *Oxford Bulletin of Economics and Statistics* (1986, 1992), the *Journal of Economic Dynamics and Control* (1988), *Advances in Econometrics* (1990) and *Econometric Reviews* (1992) for additional coverage of the field.

PETER C.B. PHILLIPS

See also ARCH MODELS; AUTOREGRESSIVE AND MOVING-AVERAGE TIME-SERIES PROCESSES; BAYESIAN INFERENCE IN TIME SERIES; COINTEGRATED ECONOMIC VARIABLES; PRESENT VALUE; RATIONAL EXPECTATIONS BUSINESS CYCLE MODELS; STATIONARITY; STATIONARY TIME SERIES; STATISTICAL INFERENCE IN TIME SERIES; STOCK PRICES AND MARTINGALES; TESTING FOR UNIT ROOTS; VOLATILITY.

BIBLIOGRAPHY

Beveridge, S. and Nelson, C.R. 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics* 7: 151–74.

Bhargava, A. 1986. On the theory of testing for unit roots in observed time series. *Review of Economic Studies* 52: 369–84.

Box, G.E.P. and Jenkins, G.M. 1976. *Time Series Analysis: Forecasting and Control.* Revised edn, San Francisco: Holden Day.

Campbell, J.Y. and Perron, P. 1991. Pitfalls and opportunities: what macroeconomists should know about unit roots. Working paper no. 100, National Bureau of Economic Research, Cambridge, Mass.

Campbell, J.Y. and Shiller, R.J. 1988. Interpreting cointegrated models. *Journal of Economic Dynamics and Control* 12: 505–22.

Christiano, L.J. 1988. Searching for a break in GNP. Working paper no. 2695. National Bureau of Economic Research, Cambridge, Mass.

DeJong, D.N. and Whiteman, C.H. 1991a. Trends and random walks in macroeconomic time series: a reconsideration based on the likelihood principle. Working paper no. 89-4, Department of Economics, University of Iowa. Forthcoming in *Journal of Monetary Economics.*

DeJong, D.N. and Whiteman, C.H. 1991b. The temporal stability of dividends and stock prices: evidence from the likelihood function. *American Economic Review* 81: 600–617.

Dickey, D.A. 1976. Estimation and hypothesis testing in nonstationary time series. Unpublished PhD dissertation, Iowa State University.