Annals Issue on Long Memory and Nonlinear Time Series: Editors' Introduction

James Davidson — Cardiff University

Timo Teräsvirta — Stockholm School of Economics

October 2001

The last two decades have witnessed tremendous advances in econometric time series research. The linear stationary framework of ARMA and VAR models driven by i.i.d. shocks, which was for many years the cornerstone of econometric modelling, has increasingly given way to methods that can deal with the manifestly nonstationary and nonlinear features of many economic and financial time series. Two types of model in particular have found their way into the mainstream of applied research, the unit-root/cointegration framework for nonstationary time series and the ARCH and related models of conditional heteroscedasticity. Recent research has been aimed at both extending our understanding of these well established models, and widening the range of data features that can be handled. Long memory models generalize the unit root model of nonstationarity, and a range of new models of nonlinear dynamics allow for asymmetric responses, threshold behaviour and stochastically switching regimes. The concept of cointegration has been generalized to accommodate many of these novel features.

The papers gathered in this special Annals issue were presented at a conference in Cardiff, UK on July 9th-11th 2000, called to bring together researchers with common interests in these topics. Thirty six invited participants from six countries took part, and a total of sixteen papers were delivered, of which twelve have been submitted and accepted, after revision, for publication in this issue. Eight of them take nonstationarity or long memory as a theme, and eight are concerned with nonlinearity. In other words, no fewer than four succeed in combining both concerns. This was not an outcome planned or anticipated by the conference organizers, but illustrates the extent to which time series research represents a thoroughly unified and interconnected enterprise.

We must take this opportunity to record our gratitude to the sponsors of the conference: the Julian Hodge Institute for Applied Macroeconomics, the Cardiff Business School, the Economic and Social Research Council's 'Understanding the Evolving Macroeconomy' Programme, and Cardiff County Council. We also thank David Peel, the co-organizer of the conference with James Davidson; Cheng Hsiao and the editorial board of the *Journal of Econometrics* for inviting us to edit this special issue; and not least, the contributors and the anonymous referees, with whose help the issue has been finalized with exceptional speed and smoothness.

Long Memory Models

Long memory, in the form of the fractional integration (FI) model, was introduced to the econometrics literature by Granger and Joyeux (1980). The fractional difference operator is defined as

$$(1-L)^{-d} = \sum_{j=0}^{\infty} \frac{d\Gamma(j+d)}{\Gamma(1+d)\Gamma(j+1)}$$

where L is the lag operator and d is a real number. The lag weights are of $O(j^{d-1})$, and hence decline to zero when $d < 1$, but are non-summable when $d > 0$, although square summable when $d < \frac{1}{2}$. Accordingly, the process

$$x_t = (1-L)^{-d} u_t$$

where u_t is a 'short memory' process such a stationary invertible ARMA(p, q), is called long memory for $d > 0$. It is covariance stationary for $d < \frac{1}{2}$, but not otherwise. When $d < 1$ the process is sometimes called 'mean reverting', although this terminology needs to be used with care since the existence of the mean is not easily shown when the variance is undefined. It is better to think of the relevant property as independence of initial conditions, since the lag weights decline eventually to 0, unlike the cases $d \geq 1$. The case $-\frac{1}{2} > d > 0$ is called anti-persistent, and in this case the lag weights sum identically to 0. This is best thought of as the simple difference of a 'mean-reverting' nonstationary process.

The interest of these processes for macro and financial econometrics is that they extend the familiar

'I(0)-I(1)' paradigm, by embedding the cases $d = 0$ and $d = 1$ in a continuum of memory properties. Although the I(1) model is remarkably good at describing many macroeconomic time series, it has some drawbacks as a model for economic data; for example, it is asymptotically unbounded in both directions, with probability 1. Even after logarithms are taken, many series observe natural bounds that contradict such a property; interest rates and unemployment rates come to mind. Note that, while it may be argued that the unit root model is adequate to describe macroeconomic data over comparatively short time spans, its key role in inference procedures is to provide an asymptotic approximation to Brownian motion. If the model is asymptotically implausible, ought this to cast doubt on the use of such approximations? Fractional integration offers one possible route to a more flexible and realistic framework for time series modelling, that can overcome this kind of critique. Examples of applied work exploiting this approach include Cheung and Lai (1993), Baillie and Bollerslev (1994), Booth and Tse (1995), Masih and Masih (1995), Sephton (1996), and Duecker and Startz (1998).

This increased generality comes at the price of some puzzles, and some apparent drawbacks. Long memory models do not possess finite-order autoregressive lag representations, and therefore cannot be generated by finite-order difference equations, whether linear or nonlinear. This poses a problem for writing down representative agent models of economic behaviour incorporating long-memory features, other than the special unit root cases, where d is an integer. However, this awkward aspect of these models may also be salutary. Granger (1981) pointed out that long memory models closely resembling FI models are induced by the aggregation of a large number of variables driven by heterogeneous difference equations. The fact that many economic variables are measured by aggregation over heterogeneous agents should not escape our notice, although little work on the theoretical aspects of fractional modelling has been done to date. See Byers, Davidson and Peel (1997, 2001) and Ding and Granger (1996) for applications of this approach to modelling, respectively, the level and the conditional variance (long memory ARCH). However, the cash value of the fractional model must ultimately lie in being able to describe observed time series and their relationships better, or more parsimoniously, than conventional difference equations. Whether it can do so is an issue to be resolved by both empirical and theoretical investigations.

The contributions to this special issue deal with a wide range of issues related to long memory modelling, with both a univariate and a multivariate emphasis. Ingolf Dittmann and Clive Granger address a simple but fundamental question; if a nonlinear transformation of a Gaussian I(d) process is taken, what are the memory properties of the transformed process? Interestingly, such transformations can make the 'memory' appear shorter in stationary cases, but not in the nonstationary cases. Their focus is on the properties of the autocorrelation functions, and hence they cannot tell the whole story of the behaviour of the transformed processes, and in particular, how successfully they might themselves be modelled as I(d). Nonetheless, their findings are often unexpected and intriguing.

Dick van Dijk, Philip Hans Franses and Richard Paap explore the interconnections between long memory and nonlinearity in a different way. They propose a model of US monthly unemployment in which the fractional differences of the series are modelled as a smooth-transition autoregressive (STAR) process. This is intended to capture the observed tendency for shocks to unemployment to follow a different regime during booms and recessions. The estimated value of d in this case is close to 0.5. This finding suggests that neither differencing the series (treating it as I(1)) nor leaving it undifferenced, before attempting nonlinear modelling, could hope to capture the memory characteristics of the series effectively.

Jorg Breitung and Uwe Hassler, and also James Davidson, consider the problem of testing for cointegrated relationships between fractionally integrated series. The former study considers tests for cointegrating rank applied to fractionally differenced series, yielding asymptotically chi-squared test criteria, while the latter proposes residual-based tests which are asymptotically non-pivotal but can be implemented using the parametric bootstrap. Both these studies address the awkward issues posed by long-memory modelling, in particular, the nature of relationships where d does not assume a 'special' value such as unity, but is apparently arbitrary. Perhaps the main virtue of these procedures is that they allow cointegration testing within a maintained hypothesis that embeds the conventional Engle-Granger (1987) setup, but is much less restrictive. It remains to be seen whether such relationships exist, and to address the major problems of interpretation and theoretical understanding they may imply. At least, a testing apparatus for this type of investigation is now in place.

Javier Hidalgo considers the constrasting problem of model selection in the context of distributed lag regressions where both regressors and disturbances are long memory but covariance stationary. The particular case of interest is the maximum lag length, where it is not ruled out that this parameter could be infinite. Such models can be estimated in the frequency domain such that the lag length may increase with sample size, but it might in practice be of low order. Hidalgo demonstrates the consistency of a selection criterion similar to that of Geweke and Meese (1981). His consistency results are also valid for commonly applied criteria such as BIC (Schwarz, 1978, Rissanen, 1978) and HQ (Hannan and Quinn, 1979).

Only two of the conference contributions relating to long memory focus on the I(1) case, although each provides an interesting twist to the standard model. Robert De Jong considers cointegration models in which the error-correction mechanism is nonlinear; for example, it might be such that the cointegrating residual is a bilinear or threshold-type process. De Jong shows that, except under additional regularity conditions, the familiar result that the distribution of the short-run adjustment parameters is the same as if the long-run (cointegrating) parameters were known may not obtain. Yoosoon Chang considers the relatively unexplored topic of testing the unit root hypothesis in panel data, where the individual cross-section units are dependent, and the panel may be unbalanced. The problem is to devise an asymptotically pivotal test statistic for the joint unit root hypothesis, not depending on nuisance covariance parameters. Chang proposes a new test based on the average of the Dickey-Fuller type t statistics where these have been computed by a nonlinear IV procedure.

Nonlinearity

Nonlinear processes, like the ones with long memory, may also extend or often blur the familiar 'I(0)-I(1)' paradigm. Series such as unemployment rates or interest rates that are often treated as I(1) processes or processes with long memory may sometimes also be viewed as generated by stationary nonlinear processes. Asymmetries in many unemployment rate series or even economic theory may suggest such an alternative; see, for example, Bianchi and Zoega (1998) or Skalin and Teräsvirta (2002). Indeed, van Dijk et al. combine nonlinearity with long memory when they model the US unemployment rate. Exchange rates fluctuating in a target zone may constitute another example. Their behaviour may be appear nonstationary well inside the zone but must be different close to the boundaries. One may also mention threshold cointegration; the strength of attraction to a linear cointegrating relationship may be zero somewhere in the sample space (local I(1)-type behaviour) and nonzero elsewhere (local I(0)-type behaviour). This is a case examined in this issue by Bruce Hansen and Byeongseon Seo, and also by Robert de Jong.

If one searches early applications of empirical nonlinear models in macroeconomics, disequilibrium models, see Fair and Jaffee (1972), may come into mind. In these models, prices do not adjust immediately with quantities, which causes nonlinearity even when the demand and supply equations are linear in parameters. A class of single-equation nonlinear models that has become popular in economics is the class of switching regression or threshold regression models. A typical member of this class may be written as

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{x}_t + \sum_{j=2}^{m+1} \boldsymbol{\beta}_j' \boldsymbol{x}_t I(z_t > s_{j-1}) + \varepsilon_t$$

where $\boldsymbol{x}_t = (1, y_{t-1}, ..., y_{t-p}, x_{1t}, ..., x_{kt})'$ is a $(k+p+1) \times 1$ vector of variables, the last k being exogenous, $\boldsymbol{\beta}_j, j = 1, ..., m+1$, are parameter vectors, $\{\varepsilon_t\}$ is a sequence of error terms, often assumed i.i.d., $I(A)$ is an indicator function: $I(A) = 1$ when the argument is true, zero otherwise, and z_t is an observable switching or threshold variable and $s_1, ..., s_m$ unknown threshold values such that $-\infty < \underline{s} < s_1 < ... < s_m < \overline{s} < \infty$. The economic relationship in question is thus characterized by a number of switches from one regression line to another. Quandt (1958) already considered this model: in his approach $z_t = t$. Bacon and Watts (1971) suggested making the transition from one regression line to the other smooth, which leads to the smooth transition regression model. Both the switching regression and smooth transition regression have found application in macroeconometric modelling. As already mentioned, the univariate nonlinear model in van Dijk et al. is a STAR model with two regimes.

Estimating switching or threshold regression models may be numerically tedious if the number of regimes or thresholds is large. This fact prompted Jesus Gonzalo and Jean-Yves Pitarakis to investigate computationally less demanding estimation algorithms, and developing one also enabled the authors to propose a bottom-up sequential modelling procedure for selecting the number of regimes in the threshold model. The key result of Gonzalo and Pitarakis is the following. Suppose that the data are generated by a threshold regression model with m thresholds but one fits a model with a single threshold. Then, under certain conditions, the (quasi) maximum likelihood estimator of this single threshold is consistent for one of the m threshold values in the original model. Another threshold value may now be estimated conditionally on the first one using the same idea, and the estimator is consistent for another threshold parameter.

The thresholds are thus estimated sequentially, which substantially alleviates the computational burden compared to simultaneous estimation using an m-dimensional grid.

The sequential estimation technique opens up a possibility of combining sequential estimation and model selection. It is possible to compute the value of a model selection criterion each time another threshold is added to the model. The number of regimes in the threshold regression may then be selected such that the value of the model selection criterion such as BIC is minimized. Gonzalo and Pitarakis discuss asymptotic properties of their selection criteria and carry out simulations with a number of selection criteria at moderate sample sizes.

The results of Gonzalo and Pitarakis concern stationary processes. Bruce Hansen and Byeongseon Seo also consider threshold regressions, but their focus is on threshold cointegration. In threshold cointegration models, cointegrating relationships between nonstationary variables have the same structure in all regimes, whereas the strength of attraction varies from one regime to the other. Hansen and Seo focus on two-regime threshold cointegration models with a single cointegrating relationship. They consider estimating the parameters of such a model and testing the null hypothesis of linearity when the cointegrating relationship is not known. The authors work out an algorithm for maximum likelihood estimators and provide asymptotic theory for the statistic they derive for testing for a threshold. The distribution theory is nonstandard as the testing situation involves the well-known problem that the model is identified under the alternative but not under the null hypothesis. Asymptotic p-values are obtained by the fixed regressor bootstrap. Another alternative, the residual bootstrap for obtaining an empirical null distribution of the test statistic, is considered as well. Both alternatives are simulated, and the test based on the residual bootstrap turns out to have better size properties than the fixed regressor bootstrap. The paper also contains an interesting application to the term structure of interest rates.

The topic of Valentina Corradi and Norman Swanson is model comparison by out-of-sample prediction. An assumption that is specific for the paper is that the models under comparison are nested. More specifically, the alternative (larger) model is a generic nonlinear model, and the same loss function is assumed in-sample (for the estimation of parameters) and out-of-sample. Testing the Granger noncausality hypothesis in a nonlinear setting may serve as an example of a situation in which such a comparison is relevant. Corradi and Swanson work out the asymptotic null distribution for the test statistic they suggest for the purpose. The distribution is nonstandard so that critical values have to be obtained by simulation. They also consider small sample properties of the test and report that they are satisfactory, at least when the loss function is quadratic.

Definitions of linear models often contain the assumption that the errors of the model are i.i.d. Models involving a nonconstant conditional variance such as ARCH and GARCH are thus nonlinear, even in cases where the error term has a constant unconditional variance. Two papers of this issue consider either GARCH or alternatives to GARCH. Joon Y. Park discusses a family of models for conditional heteroskedasticity in which the heteroskedasticity is controlled by a nonstationary integrated variable. These processes are said to contain nonstationary nonlinear heteroskedasticity (NNH). The NNH models are capable of generating volatility clustering and produce series with leptokurtosis. The statistical properties of the model depend on heterogeneity-generating functions, and two types of such functions, previously considered by Park and Phillips (1999, 2001) are used in the paper. Estimation of parameters can be carried out by nonlinear least squares. The NNH model is applied to a weekly USD/DEM exchange rate series and the results compared to those obtained from standard ARCH(1) and GARCH(1,1) models.

Stefan Lundbergh and Timo Teräsvirta observe that GARCH modellers do not often test the adequacy of their estimated model, although this is quite common when modelling the conditional mean. They derive a number of simple misspecification tests for standard GARCH models and compare them with existing tests. They test the null hypothesis of no remaining ARCH, linearity, and parameter constancy, the alternative in the last-mentioned test being smoothly changing parameters. The tests are Lagrange multiplier (LM) or LM type tests. In simulations, their power compares favourably with the power of the tests hitherto considered in the literature.

Concluding Remarks

Given a larger conference budget, not to mention unlimited space in this Journal, there are a number of other important research topics that we would like to have seen addressed here. Markov-switching models of the conditional mean and/or the conditional variance are a class of models now receiving a lot of attention; for a recent application to the unit root testing problem, for example, see Nelson, Piger and Zivot

(2001). Long memory models of conditional heteroscedasticity such as that of Ding and Granger (1996) represent yet another fruitful blending of themes. Another topic attracting a lot of interest currently is the link between long memory and nonlinearity in mean, specifically, the ability of certain nonlinear models incorporating random switches of regime to yield autocovariance structures very similar to the (linear) fractional model; see for example Diebold and Inoue (2001). Such omissions notwithstanding, the present collection of papers appears to us to give a fairly representative picture of the work going on currently in econometric time series research.

References

Bacon, D.W. and D.G. Watts, 1971, Estimating the transition between two intersecting straight lines, Biometrika 58, 525-534.

Baillie, R. T. and Bollerslev, T., 1994, Cointegration, fractional integration and exchange rate dynamics, Journal of Finance 49, 737–45.

Bianchi, M. and Zoega, G., 1998, Unemployment persistence: does the size of the shock matter? Journal of Applied Econometrics 13, 283-304.

Booth, G. B. and Tse, Y., 1995, Long memory in interest rate futures markets: a fractional cointegration analysis, Journal of Futures Markets 15, 573–84.

Byers, D., Davidson, J. and Peel, D., 1997, Modelling political popularity: an analysis of long-range dependence in opinion poll series, Journal of the Royal Statistical Society (A) 160, 471-90.

Byers, D., Davidson, J. and Peel, D., 2001, Modelling political popularity: a correction, Journal of the Royal Statistical Society (A) (forthcoming).

Cheung, Y.-W. and Lai, K. S., 1993, A fractional cointegration analysis of purchasing power parity, Journal of Business and Economic Statistics 11, 103–12.

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

Ding, Z. , and Granger, C.W.J., 1996, Modelling volatility persistence of speculative returns: a new approach, Journal of Econometrics 73, 185-215.

Duecker, M. and Startz, R., 1998, Maximum likelihood estimation of fractional cointegration with an application to US and Canadian bond rates, Review of Economics and Statistics 80, 420–6.

Engle, R. F. and Granger, C.W.J., 1987, Cointegration and error correction: representation, estimation and testing, Econometrica 55, 251-75.

Fair, R.C. and D.M. Jaffee, 1972, Methods of estimation for markets in disequilibrium, Econometrica 40, 497-514.

Granger, C. W. J., 1981, Long memory relationships and the aggregation of dynamic models, Journal of Econometrics 14, 227-38.

Granger, C. W. J. and Joyeux, R., 1980, An introduction to long memory time series models and fractional differencing, Journal of Time Series Analysis 1, 15–29.

Hannan, E. J. and B. G. Quinn, 1979, The determination of the order of an autoregression. Journal of the Royal Statistical Society (B) 41, 190–5.

Masih, R and Masih, A. M. M., 1995, A fractional cointegration approach to empirical tests of PPP: new evidence and methodological implications from an application to the Taiwan/US Dollar relationship, Weltwirtschaftliches Archiv 131, 673–94.

Nelson, C. R.,J. Piger and E. Zivot, 2001, Markov regime switching and unit-root tests, Journal of Business & Economic Statistics, 19, 4, 404-415.

Quandt, R., 1958, The estimation of the parameters of a linear regression obeying two separate regimes, Journal of the American Statistical Association 53, 873-880.

Park, J.Y. and P.C.B. Phillips, 1999, Asymptotics for nonlinear transformations of integrated time series, Econometric Theory 15, 269-98.

Park, J.Y. and P.C.B. Phillips, 2001, Nonlinear regressions with integrated time series, Econometrica 69, 117-61.

Rissanen, J., Modeling by shortest data description, Automatica 14, 465-71.

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

Sephton, P. S., 1996, A note on fractional cointegration, Applied Economics Letters 3, 683–5.

Skalin, J. and T. Teräsvirta, 2002, Modelling asymmetries and moving equilibria in unemployment rates, Macroeconomic Dynamics 6, forthcoming.