Cowles Foundation Reprint 885

Econometric Theory, 10, 1994, 453-460. Printed in the United States of America.

BAYES METHODS AND UNIT ROOTS

Editors' Introduction

PETER C.B. PHILLIPS
Cowles Foundation for Research in Economics
Yale University

HERMAN K. VAN DIJK
Econometric Institute and Tinbergen Institüt
Erasmus University Rotterdam

1. INTRODUCTION

Since the mid 1980's there has been enormous growth of interest in the statistical analysis of trending time series. Recent research has focused on methods that attempt to distinguish the nature of the trending mechanism and, in particular, to determine whether it is best modeled through deterministic time trends (such as polynomial functions of time) or stochastic trends (such as random walks or integrated processes). A concomitant interest in the applied macroeconometric literature has been the decomposition of aggregate time series such as GNP into components that may be deemed persistent and those that may be deemed transitory. Such decompositions are important in macroeconomic thinking and policy analyses because the implications of persistent shocks are quite different from those that are transitory.

Econometricians and statisticians are also interested in the effects of stochastically trending time series on forecasting. The practical implications for forecasting the long-run behavior of a set of trend-stationary series are very different from those of time series with some unit roots and possibly some cointegration. The effects of nonstationarity in forecasting are particularly relevant when one has to evaluate the long-term effects of investments. For example, in evaluating public investments in economic infrastructure like information highways, high-speed trains, or large airports it is necessary to forecast the effects of the investments over time horizons such as 15-20 years. How we model the nonstationarity in the data then has a big effect on forecasts and forecast-confidence intervals over such long time horizons. Similar considerations apply to insurance companies in determining the level of pension premiums that are designed to cover indexed pensions in 20-30 years' time.

 0266-4666/94 $5.00 + .00

Although much of the emerging literature on the topic of nonstationary time series has been written by econometricians with economic applications in mind, many of the results and methods are of more general statistical interest and have been obtained by statisticians and probabilists. Indeed, initial work by Fuller [6] and Dickey and Fuller [3,4] on statistical testing for Gaussian random walks provided a starting point for subsequent work. Solo [21], Phillips [12], and Chan and Wei [1] demonstrated the utility of functional limit theory in this context and, most recently, Jeganathan [8] employed these methods to develop a general asymptotic theory of inference in models that permit roots on the unit circle. Thus, the field has provided a very active interface between statistics and econometrics.

There are many subjects of ongoing interest and development that seem appropriate for a fertile exchange of ideas in this field. Foremost among these at the present time is the question of Bayesian versus classical methods of inference concerning trends in time series. This is a topic on which there is now active research [2,13,17,19,20] and yet there are still many issues to be resolved, including the logical formulation of priors that properly accommodate nonstationary time series models, Bayesian modeling that permits general forms of weak dependence in time series, Bayesian asymptotics that allow for nonstationary processes, Bayesian treatment of non-Gaussian data, and computational issues. Also important, but only now under development [5,14] is a classical theory of optimality in testing for the presence of unit roots and cointegration.

This special-themed double-issue of *Econometric Theory* draws together researchers from statistics, probability, and econometrics to address some of these topics and provides a productive exchange of perspectives between Bayesian and classical statistical methodologists.

2. THE YALE-NSF CONFERENCE SERIES

The papers included in this issue (with three exceptions) were all presented at a conference on "Bayes Methods and Unit Roots" that was held at Yale in the spring of 1992. All of the papers included are written by conference participants. The conference was the first in a series at Yale on the general theme of "Applications of Functional Limit Theory to Econometrics and Statistics." The conference series, which at the time of writing is still ongoing, is supported by a grant to Yale from the National Science Foundation, and its intention is to foster the growing interaction between professional statisticians and econometricians that is taking place in this general subject area.

The conference on "Bayes Methods and Unit Roots" was held over a two-day period in April 1992 and was jointly organized by Peter Phillips and Christopher Sims. The final program of this conference is printed at the end of this series of papers. Twenty-three visitors from outside of Yale attended the conference, which contained three invited lectures delivered by James

Berger, Herman van Dijk, and Bruce Hill, and twelve contributed papers. In preparing this symposium double-issue of *ET*, all of the papers have followed our normal review process, and we thank the authors, conference participants, and external referees for their assistance in making possible this joint publication of the conference proceedings.

3. CONTENTS OF THIS SYMPOSIUM ISSUE

The papers brought together here are the direct response to our invitation to contribute to the general subject of "Bayes Methods and Unit Roots." We did not lay down any preconceived notion of which topics were the most relevant or important in this general subject area when we approached potential participants (although the original Yale proposal to the NFS did lay out interesting ongoing themes as described in the Introduction). The grass-roots response that we received from our contributors gave rise to the following topics:

- the choice of a noninformative prior in the univariate and the multivariate autoregressive model (Berger and Yang; Kleibergen and van Dijk; Uhlig; Schotman);
- the importance of model parameterization, the choice of priors, and the effect of initial values on posterior and predictive distributions (Kleibergen and van Dijk; Schotman; Uhlig; Zivot);
- the effect of thick-tailed rather than Gaussian distributions on Bayes inference (Geweke);
- the use of Markov switching principles to entertain dual specifications (trend- and difference-stationarity) for each observation (McCulloch and Tsay);
- data-based model selection with posterior odds for possibly nonstationary time series (Phillips and Ploberger);
- Bayesian forecasting for economic time series (Hill);
- classical pretest problems and Bayesian alternatives for inference when the order of integration is unknown (Elliot and Stock);
- testing the null hypothesis of stationarity (Choi);
- Bayesian encompassing tests and unit roots (Florens, Larribeau, and Mouchart).

Berger and Yang study and compare various approaches to the construction of a noninformative prior for the AR(1) model. They show that the "reference prior" approach works satisfactorily for the stationary case but not for the explosive case. To address the difficulties that are encountered in the explosive region of the parameter space, they develop and recommend a symmetric version of the stationary reference prior. Simulations show that the symmetrized reference prior works reasonably well in terms of the implied sampling properties of the posterior mean and Bayes confidence sets.

Hill puts forward a modeling procedure that is based on Bayesian principles for forecasting economic time series. The idea is that future observations can be predicted by taking a weighted average of the optimal predictions of two competing models in which the weights are delivered by the Bayesian posterior probabilities of the models. An illustration is provided where one model is a simple random walk and the alternative is an external shock (or significant real-world event) that frees the data from its past behavior. This approach to forecasting is related to the multiprocess modeling ideas suggested by Harrison and Stevens [7] and exposited in West and Harrison [22]. In effect, the predictions are generated by averaging over models and thereby incorporate different ways of representing the past data.

Kleibergen and van Dijk introduce a cointegration model in which cointegration may be tested by a variable addition procedure. By using flat priors, they show that the marginal posteriors of the cointegration vectors are ill-behaved when certain parameters become nonidentified, which occurs when the model is difference-stationary. They propose the Jeffreys prior, which is proportional to the square root of the determinant of the information matrix, and discuss different cases of the Jeffreys prior. Special attention is given to the effect of the initial values in a multivariate autoregressive framework. Kleibergen and van Dijk's methods need to be applied to other economic examples to assess the empirical value of their work.

Zivot considers the unobserved components representation of an autoregressive model. The issue here is that in the presence of a unit root component the location parameter of a deterministic trend is no longer identified. The use of improper priors yields improper posteriors in such cases. By analyzing carefully the effect of the initial observation, Zivot presents a solution to this difficulty in which the posteriors are proper and posterior odds analysis is possible.

Schotman studies the effect of different parameterizations and prior dependence between parameters in an AR(1) model with an intercept or supplementary regressor. By working with an error correction model formulation, Schotman gives a prior density whose limiting forms (as certain parameters approach the limits of their domain of definition) include the uniform and the Jeffreys density. Schotman shows that the posterior density can be very sensitive to the degree of prior dependence between the parameters. In the case of an AR(1) with an intercept or an unconditional mean, the weaker the prior dependence between the mean and the autoregressive coefficient, the more the posterior of the autoregressive parameter is shifted toward a unit root. As in the case of the study by Zivot, this conclusion is affected by the local identification problem that occurs for the intercept when the AR parameter is unity.

McCulloch and Tsay treat trend- and difference-stationary models as two competing hypotheses for each observation and introduce a Markov switching scheme to estimate the probability of the appropriate state for each observation. A Gibbs sampling procedure is introduced as a computational procedure. The proposed method allows the researcher to monitor the evolution of the two competing hypotheses over time. An empirical illustration of the technique shows that postwar U.S. monthly industrial production is generally well modeled as difference-stationary except for the latter part of the series where the classification is less clear.

Geweke relaxes the assumption of normally distributed disturbances to the case of Student-t distributed disturbances in an autoregressive model representation, where the prior density on the autoregressive parameter depends on the sampling interval of the time series. By using a Gibbs sampling technique, Geweke obtains posterior results of the parameters of interest. His empirical findings suggest that the move from normally distributed to Student-t distributed disturbances is important in many cases. This confirms parallel research of Kleibergen and van Dijk [9] who concluded that the step from homoskedastic to heteroskedastic disturbances is less important than the step from normal to Student-t disturbances.

Uhlig calculates the Jeffreys prior for an AR(1) process by using the exact likelihood. The latter is computed under the assumption that the process started up at some distant but finite date from the first recorded observation (y_0). Uhlig compares the form and properties of Jeffreys prior constructed from this schema with the Jeffreys prior that is derived from the data-density conditional on y_0, as in Phillips [13]. His results in this analysis support the use of a flat prior in the nonexplosive region, as used in Sims and Uhlig [19] only when $y_0 = 0$ and there is no constant or trend. In other cases, there are major differences between the Jeffreys prior and a flat prior.

In a second paper, Uhlig gives a personal view and discussion from a Bayesian perspective of some rules he argues are useful for applied researchers and macroeconomists who are analyzing nonstationary economic time series. Next, he discusses the consequences of the presence of unit roots for medium-term forecasting, again from a Bayesian point of view. By taking parameter uncertainty into account, he obtains predictive distributions that are asymmetric and have tails that are sensitive to the prior treatment of explosive roots.

Elliot and Stock consider a bivariate regression model where the regressor may be $I(0)$ or $I(1)$. Their focus of interest is inference on the regression coefficient, and they show the size distortions that can arise from either ignoring potential nonstationarity or using unit root pretest procedures to determine the appropriate asymptotics. They propose an alternative mixture approximation to the limit distribution of the usual regression t-test that is a mixture of the two conditional asymptotic distributions corresponding to the $I(0)$ and $I(1)$ possibilities. The mixture itself relies on a statistic that effectively selects the class of the regressor as the sample size gets large and thereby has the correct asymptotic size.

DeJong and Whiteman conduct a Bayesian analysis of the validity of the restrictions that the present-value model imposes on a vector autoregressive model. They compute a weighted average of the conditional probability of restrictions given trend- and difference-stationarity where the weights are given by the prior probabilities of the two states. Given a tight prior, which is popular in forecasting, the present-value model is not favored by the data, while allowing for more prior uncertainty yields a more favorable outcome for the present value model.

Choi develops classical tests for the null hypothesis of level- and trend-stationarity by using the LM principle. The idea is similar to the LM test for stationarity in Kwiatkowski et al. [10] and the LM test for an MA unit root in Saikkonen and Luukkonen [15]. The limit distributions of the tests are nonstandard functionals of Brownian motion, but Choi gives alternative exact expressions for them in terms of chi-square statistics and the Gaussian error function. Application of this procedure and conventional unit root tests to U.S. macroeconomic time series gives mixed results: for some series, the inferences concerning stationarity unit roots are compatible; for many series they are not, indicating that the discriminatory power of the data is not strong, at least through the medium of these tests.

Florens, Larribeau, and Mouchart illustrate the Bayesian approach to encompassing by using the simple Gaussian AR(1) model with a known variance. The model with and without a unit root is considered, and an encompassing procedure is mounted to test whether the root is unity. The procedure involves the computation of a Bayesian pseudotrue value, which permits the extension of the null model (here the Gaussian random walk) to the alternative (the Gaussian AR(1)). Inference on the "autoregressive parameter" is then possible within the extended null model and the posterior distribution so derived can be compared with that from the alternative model. The authors provide details of the numerical procedures, including simulation-based integrations, that are needed to execute the steps in this procedure.

Kim studies an AR(1) model driven by errors that are not necessarily Gaussian. He shows that the posterior distribution of the autoregressive coefficient is asymptotically normal under quite general conditions that allow for the presence of a unit root in the true data-generating process. This result is important to practitioners because it means that large sample Bayesian inference may be conducted in a consistent way irrespective of the presence of a unit root in the generating mechanism. Unlike classical asymptotics, nonstandard limit distributions are not required. Kim's work is related to other recent work on this topic by Sims [18] and Phillips and Ploberger [14].

Phillips and Ploberger develop Bayesian inference and model-selection procedures for the stochastic linear regression model with Gaussian errors. Under a uniform prior, the Bayesian data density is shown to belong to the exponential family. Even though this measure is improper, martingale properties still apply and, in particular, both the maximum likelihood estimator and the posterior distribution are local martingales under this measure. The measure also leads to a new model-selection criterion called "PIC" (a posterior information criterion), which allows for nonstationary data. PIC is asymptotically equivalent in stationary systems to the commonly used Schwarz [16] criterion BIC. The theory is applied to ARMA models with deterministic trends and used for automated order selection of the stochastic regressors and the trend degree. Simulations show that the procedure works well in practice both for nonstationary and stationary systems. An empirical implementation of these methods to the Nelson–Plosser [11] series is given to illustrate their use in practice. Only three of the series (unemployment, industrial production, and the money stock) are found to be level or trend-stationary and the remaining series are found to have unit roots.

REFERENCES

1. Chan, N.H. & C.Z. Wei. Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16 (1988): 367–401.
2. DeJong, D.N. & C.H. Whiteman. Reconsidering trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 28 (1991): 221–254.
3. Dickey, D.A. & W.A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74 (1979): 427–431.
4. Dickey, D.A. & W.A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49 (1981): 1057–1072.
5. Elliott, G., T. Rothenberg & J. Stock. Efficient Tests for an Autoregressive Unit Root. Mimeo, 1992.
6. Fuller, W.A. *Introduction to Statistical Time Series*. New York: Wiley, 1976.
7. Harrison, P.J. & C.F. Stevens. Bayesian forecasting (with discussion). *Journal of the Royal Statistical Society, Series B* 38 (1976): 205–247.
8. Jeganathan, P. On the asymptotic behavior of least squares estimators in AR time series with roots near the unit circle. *Econometric Theory* 7 (1991): 269–306.
9. Kleibergen, F. & H.K. van Dijk. Non-stationarity in GARCH models: A Bayesian analysis. *Journal of Applied Econometrics* 8 (1993): S41–S61.
10. Kwiatkowski, D., P.C.B. Phillips, P. Schmidt & Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54 (1992): 159–179.
11. Nelson, C.R. & C. Plosser. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics* 10 (1982): 139–162.
12. Phillips, P.C.B. Time series regression with a unit root. *Econometrica* 55 (1987): 277–301.
13. Phillips, P.C.B. To criticize the critics: An objective Bayesian analysis of stochastic methods. *Journal of Applied Econometrics* 6 (1991): 333–364.
14. Phillips, P.C.B. & W. Ploberger. Time Series Modeling with a Bayesian Frame of Reference: Concepts, Illustrations and Asymptotics. Cowles Foundation discussion paper, Yale University, rev., 1992. (original draft, 1991)
15. Saikkonen, P. & R. Luukkonen. Testing for a moving average unit root in autoregressive moving average models. *Journal of the American Statistical Association* 88 (1993): 596–601.
16. Schwarz, G. Estimating the dimension of a model. *Annals of Statistics* 6 (1978): 461–464.
17. Sims, C.A. Bayesian skepticism on unit root econometrics. *Journal of Economic Dynamics and Control* 12 (1988): 436–474.

18. Sims, C.A. Asymptotic Behavior of the Likelihood Function in an Autoregression with a Unit Root. Mimeo, Yale University, 1990.
19. Sims, C.A. & H. Uhlig. Understanding unit rooters: A helicopter tour. *Econometrica* 59 (1991): 1591-1600.
20. Schotman, P.C. & H.K. van Dijk. A Bayesian analysis of the unit root in real exchange rates. *Journal of Econometrics* 49 (1991): 195-238.
21. Solo, V. The order of differencing in ARIMA models. *Journal of the American Statistical Association* 79 (1984): 916-921.
22. West, M. & P.J. Harrison. *Bayesian Forecasting and Dynamic Models*. New York: Springer-Verlag, 1989.