A. Banerjee

Structural breaks have been the focus of much recent attention in the econometrics literature, see e.g. the special issues of the Journal of Business and Economic Statistics (1992) and the Journal of Econometrics (1996) devoted to this topic. Starting at least with the work of Durbin and his co-authors (see e.g. Brown, Durbin and Evans (1975)) on CUSUM-type statistics, research into tests of stability of regression models has remained a live, useful and necessary topic for discussion in both theoretical and applied econometrics.

In the jargon of Engle, Hendry and Richard (1983), the stability of an estimated model is crucial to the task of evaluating the impact of policy changes within a macroeconomic system. If a model is not invariant to policy interventions, no meaningful discussion of the kind "What would happen if ....", could be undertaken since the Lucas (1976) critique would apply with full force and policy simulation analyses would become extremely problematic.

The challenge of applied econometric modelling is therefore to develop data-congruent models which are invariant to interventions. Thus, say, one models the conditional density of y, conditional on variables z and w, and the marginal processes generating z and w change, for policy analyses to be undertaken without difficulty one would need to have stability of the conditional density. Testing of (for) the Lucas critique (Ericsson and Irons (1995)) consequently becomes a significant part of the evaluation of large and small macroeconomic models, and the devising of powerful tests of stability thereby becomes a vitally important task for the theoretical econometrician.

The development of structural stability tests has been undertaken in several related contexts and providing a unifying framework for all the various problems tackled remains a challenging issue. Two broad strands can nevetheless be identified. The first takes as its starting point the null hypothesis of structural stability, by which is meant that the data generating process remains unaltered over the period of time under investigation. The form of the alternative hypothesis is usually left fairly general, to incorporate changes in any arbitrary interesting subset of parameters of the process. These could therefore include changes in the mean, or the trend growth rate of the process, or the values of the coefficients on the independent variables (say, in a linear regression model) or in the variance of the error process. Tests would have more or less power, depending on their construction, against various classes of alternative hypotheses. The paper by Durban and Evans cited above would fall in this category as would more recent papers by Andrews, Ploberger, Kramer, with various co-authors (for further information, see references in Banerjee, Lumsdaine, Stock (1992) and in the special issues of the JBES and JOE cited above).

The second strand takes as its starting point the paper by Perron (1989) and parameterises the form of the deviation from the null hypothesis much more precisely. Perron's discussion focuses on a long data set which includes the Great Depression and the 1973 Oil Shock and the departures from stability considered are changes in the mean and/or trend growth rates of the macro-variables such as GNP, Real Wages, Industrial Production or Employment. Thus under this hypothesis a series y_t could be generated by

(1) $y_t = \mu + \beta D_t^1 + \varepsilon_t$

D_t^1 ,where is an indicator function of the form **I**(t > k) and k is the time the break occurs under the alternative, with the full sample running from 1 to T. Under the null, β is of course equal to zero. A similar parameterisation could be undertaken to denote changes in the trend of this simple function.

The second strand of argument admittedly narrows the scope of the discussion somewhat but is nevertheless significant for two reasons. First, and the original motivation for Perron's paper, is the confusion which arises in distinguishing a process of the form given by (1), for non-zero (, from a random walk and second, because such simple dummy variables for constant or trend are very often used in estimated econometric models to stabilise them. Given the richness of economic

interventions, some economic rationale is usually discernible for almost any choice of time period for the dummy variable.

Most of the theoretical discussion within the second strand concentrates somewhat frustratingly on a single-equation framework while most of the interesting empirical applications arise much more naturally within the systems framework (especially if one considers large-scale macro-modelling). Moreover, the mathematics are sufficiently complex (especially in the case where the variables are integrated) not to allow a rich class of alternatives (even if parameterised more precisely than the first class of papers) to be considered. Thus (1) has been extended but usually only within the single equation framework and only in the context of static models (ie. where all the variable are current-dated). A further useful but possibly unrealistic simplification that has been adopted is to take the timing of the break as known. Moreover the number of changepoints is usually also assumed to be known. Both of these latter simplifications are important topics for philosophical and theoretical debate with the literature.

The following agenda for discussion can be easily established:

1. How should one approach the problem within the context of empirical modelling? Which strand of the literature is more relevant when dealing with single equations or systems of equations, in the context of judging the stability of estimated models, error-correction mechanisms or co-integrating equations? Should the range of alternatives be left unspecified?

2. Leading on from above, is stabilising via dummy variables a good idea? Would the Lucas critique fail, if after such stabilising the coefficients of the original model (i.e. prior to the policy intervention) are recovered or is the very need for stabilising tools evidence that the Lucas critique holds in practice? Is the development of structurally stable models across all possible interventions or changes in parameters of the marginal and conditional processes a realistic goal?

3. How does one develop good tests which not only (a) allow for greater generality by incorporating systems of equations where each equation of the system is broken at potentially different points in time but also (b) endogenise the timing of the breaks and (c) the number of breaks? Surely this is the most relevant but, unsurprisingly, the most difficult case to deal with.

I have some limited progress to report on this front (for a narrow range of cases) dealing with (a) and (b) but the theoretical consistency and distribution properties are ill-established and need much greater investigation. What testing principle (WALD, LR or LM) might be most suitable in terms of power properties and tractability?

4. More generally, is this a useful way to look at the problem or does one miss the wood for the trees? Is there a more meaningful and tractable way of dealing with these issues? One could develop models which are structurally stable but which have no economic content. A slightly outdated line of this sort of argument would propose developing different models for different purposes, with or without economic or structural content. I would not, I think, wish to pursue this line but it might hold favour with some ways of proceeding.

5. In relation to 4., interdisciplinary approaches are likely to be exceedingly valuable. We have tried to be cautious in looking up the literature in probability and statistics, where almost all the original tools were developed, but even the most assiduous search could not reveal all that is relevant and the input of an expert in the relevant field would be of great value.

The claim, finally, is that the testing for structural stability has a rich and interesting history in the econometrics literature and continues to attract much useful and advanced work. That it has strong practical relevance is clearly beyond doubt. Nor is there much doubt that much more work remains to be accomplished. The econometricians have merely nibbled at the edges of a vastly complex cake and with the help of mathematicians and probability theorists, there is much else that could be achieved.

SELECTED BIBLIOGRAPHY

[1] Banerjee, A., Lumsdaine, R.L. and Stock, J.H. (1992). "Recursive and sequential tests of the unit - root and trend-break hypotheses: Theory and International Evidence", Journal of Business and Economic Statistics, 10, 271-288

[2] Brown, R.L., Durbin, J. and Evans, J.M. (1975). "Techniques for testing for the constancy of regression relationships over time (with discussion)", Journal of the Royal Statistical Society B, 37, 149-192.

[3] Engle, R.F., Hendry, D.F. and Richard, J-F. (1983). "Exogeneity", Econometrica, 51, 277-304.

[4] Ericsson, N.R. and Irons, J.S. (1995). "The Lucas critique in practice: Theory without measurement", in Hoover, K.D. (ed.), Macroeconometrics: Developments, Tensions and Prospects. Dordrecht:Kluwer Academic Press.

[5] Lucas, R.E. (1976). "Econometric policy evaluation: A critique", in Brunner, K. and Meltzer,A. (eds.), The Phillips Curve and Labor Markets, Vol. 1 of Carnegie-Rochester Conferences on Public Policy, pp. 19-46. Amsterdam: North-Holland Publishing Company.

[6] Perron, P. (1989). "The great crash, the oil price shock and the unit root hypothesis", Econometrica, 57, 1357-1361.