THE UNIVERSITY OF HULL

BAYESIAN MULTIVARIATE TIME SERIES MODELS FOR FORECASTING EUROPEAN MACROECONOMIC SERIES

being a Thesis submitted for the Degree of

Doctor of Philosophy

in the University of Hull

by

QIANG FU

MA (University of Manchester)
BSc (Guilin Institute of Electronic Engineering)

April, 2000

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

To my family

CONTENTS

ACKNOWLEDGEMENTS

I am very grateful to my supervisor, Professor A. R. Tremayne, for his extremely valuable, perceptive and efficient supervision and direction. I thank him not only for his careful and substantial corrections and comments, sentence by sentence, and page by page, but also for his reasons in making all those corrections. His supervision was invaluable in helping me to cope with various research problems and the final production of this thesis would have been impossible without his indispensable guidance.

I am also very grateful to Dr. Christopher Tsoukis for his advice and unstinted help.

In addition, a number of people have helped directly or indirectly with the preparation of the manuscript. Particularly, a special gratitude is due to D.M. Woodward, C.J. Hammond, J.P. Atkins, B. Williams, K. Spry and H. Stone. My fellow Ph.D. colleagues M.A. Nolan, N. Azimi, G. Agiomirgianakis, A.H. Alyousha are thanked for many helpful suggestions, stimulating discussions and constructive comments. Moreover, I greatly appreciate the social support that I enjoyed at the Department of Economics, School of Economic Studies, University of Hull whilst this thesis was formulated and written.

I am greatly appreciative of the generous financial support I have received for my doctoral research over a three-year period from the University of Hull Award and the Overseas Research Students (ORS) Award at the University of Hull.

Last, but far from least, my thanks go to my wife, Yundi, and my sister, Yu, for bearing with me and sustaining me during the difficult stretches of work on the thesis. I would also like to thank my parents and my lovely nine-year-old daughter, Lili, for their blessings, sent all the way from China.

Of course, responsibility for any errors that unfortunately remain in this thesis is entirely my own.

ABSTRACT

Research on and debate about 'wise use' of explicitly Bayesian forecasting procedures has been widespread and often heated. This situation has come about partly in response to the dissatisfaction with the poor forecasting performance of conventional methods and partly in view of the development of computational capacity and macro-data availability. Experience with Bayesian econometric forecasting schemes is still rather limited, but it seems to be an attractive alternative to subjectively adjusted statistical models [see, for example, Phillips (1995a), Todd (1984) and West & Harrison (1989)]. It provides effective standards of forecasting performance and has demonstrated success in forecasting macroeconomic variables. Therefore, there would seem a case for seeking some additional insights into the important role of such methods in achieving objectives within the macroeconomics profession.

The primary concerns of this study, motivated by the apparent deterioration of mainstream macroeconometric forecasts of the world economy in recent years [Wallis (1989), pp.34-43], are threefold. The first is to formalize a thorough, yet simple, methodological framework for empirical macroeconometric modelling in a Bayesian spirit. The second is to investigate whether improved forecasting accuracy is feasible within a European-based multicountry context. This is conducted with particular emphasis on the construction and implementation of Bayesian vector autoregressive (BVAR) models that incorporate both a priori and cointegration restrictions. The third is to extend the approach and apply it to the joint-modelling of system-wide interactions amongst national economies. The intention is to attempt to generate more accurate answers to a variety of practical questions about the future path towards a united Europe.

The use of BVARs has advanced considerably. In particular, the value of joint-modelling with time-varying parameters and much more sophisticated prior distributions has been stressed in the econometric methodology literature. See e.g. Doan *et al.* (1984),

Kadiyala and Karlsson (1993, 1997), Litterman (1986a), and Phillips (1995a, 1995b). Although trade-linked multicountry macroeconomic models may not be able to clarify all the structural and finer economic characteristics of each economy, they do provide a flexible and adaptable framework for analysis of global economic issues.

In this thesis, the forecasting record for the main European countries is examined using the 'post mortem' of IMF, OECD and EEC sources. The formulation, estimation and selection of BVAR forecasting models, carried out using Microfit, MicroTSP, PcGive and RATS packages, are reported. Practical applications of BVAR models especially address the issues as to whether combinations of forecasts explicitly outperform the forecasts of a single model, and whether the recent failures of multicountry forecasts can be attributed to an increase in the 'internal volatility' of the world economic environment. See Artis and Holly (1992), and Barrell and Pain (1992, p.3).

The research undertaken consolidates existing empirical and theoretical knowledge of BVAR modelling. It provides a unified coverage of economic forecasting applications and develops a common, effective and progressive methodology for the European economies. The empirical results reflect that in simulated 'out-of-sample' forecasting performances, the gains in forecast accuracy from imposing prior and long-run constraints are statistically significant, especially for small estimation sample sizes and long forecast horizons.

KEY WORDS AND PHRASES

Bayesian vector autoregression (BVAR) models; Exogeneity; Causality; Nonstationarity; Unit roots; Regime shifts; Common stochastic trends (CSTs); Cointegration; Error-correction (EC); Model selection; Lucas critique; Hypothesis testing; Minnesota prior; Theil U statistic.

LIST OF ABBREVIATIONS

ADF	augmented Dickey-Fuller
AEG	augmented Engle-Granger
AIC	Akaike information criterion
AR	autoregressive (process)
BIC	Bayesian information criterion
BVAR	Bayesian vector autoregressive (process)
CPI	consumer price index
CST(s)	common stochastic trend(s)
CV(s)	cointegrating vector(s)
DF	Dickey-Fuller
DGP	data generation process
DS	difference stationary
DSEM	dynamic simultaneous equations model
DW	Durbin-Watson
EC	error correction
ECM(s)	error correction mechanism(s)
FIML	full information maximum likelihood
FPE	final prediction error (criterion)
GLS	generalised least squares
GM(s)	generation mechanism(s)
HQ	Hannan-Quinn (criterion)
IFS	International Financial Statistics
IID	independently and identically distributed
IMF	International Monetary Fund
KF	Kalman filter
LM	Lagrange multiplier
LR	likelihood ratio
LS	least squares
MA	moving average (process)
MAPE	mean absolute percentage error
MCM	multi-country modelling
MDM	mainstream data-mining
ME	mean error
MEI	main economic indicators

MIDAS	Manchester Information Datasets and Associated Services
ML	maximum likelihood
MTS	macroeconomic time series
NID	normally and independently distributed
OLS	ordinary least squares
pdf	probability density function
REH	rational expectations hypothesis
RMSE	root mean squared error
SEE	standard error of estimate
SEM	structural econometric model
SSR	sum of squared residuals
SUR	seemingly unrelated regression
TBR	treasury bill rate
TS	trend stationary
UVAR	unrestricted vector autoregressive (process)
VAR	vector autoregressive (process)
VAR-ECM	vector autoregressive error correction model
VARMA	vector autoregressive moving average (process)

LIST OF NOTATION

$\equiv$	identically equals
$\Rightarrow$	implies
$\rightarrow$	converges to
$\sim$	is distributed as
$\wedge$	intersection of two planes in the space
L	lag operator
∇	first-difference operator
$\otimes$	Kronecker product or direct matrix product
$\in$	element of; as a rule, elements of a set will be denoted by the same letter as the set
$\subset$	(not necessarily proper) subset of
$\cap$	intersection of two sets, i.e. collection of elements that belong to both sets
$\cup$	union of two sets, i.e. collection of elements that belong to either set
K	dimension of a stochastic process
p	VAR order
r	cointegration rank
T	sample size
$\lvert\xi\rvert$	absolute value or modulus of ξ
diag(•)	diagonal-element matrix
exp(•)	exponential function
l(•)	likelihood function
p(•)	probability distribution function
rk($\underline{X}$)	rank of $\underline{X}$
tr(X)	trace of $\underline{X}$
vec(•)	column stacking operator

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO MACROECONOMIC MODELLING USING BAYESIAN VECTOR AUTOREGRESSIONS

§1.1 Introduction

The inherent complexity of the structure of the world economy and the increasing awareness of the need for closer policy coordination among European countries have together resulted in a growing demand for powerful methods to interpret this complicated universe and hence to strengthen the European edifice. There can be no doubt that the more accurate and consistent are ex ante forecasts and macro policy studies, the more reliable and useful will be the macroeconometric models constructed. In consequence, the resulting forecasts and simulations are likely to be more important and influential to the private, public and overseas sectors. Moreover, macroeconometric forecasting models present not only a forum within which macroeconomic theories encounter extensive challenge, but also a target upon which major controversies are likely to focus.

The importance of empirical econometric modelling naturally raises certain questions. What are the inevitable basic assumptions, or postulates, underpinning an econometric model or forecasting venture? What are the central problems associated with inferential procedures? And what should be the reaction if such problems occur? For a long time it has been widely held that macroeconometric models can be viewed primarily as providing us with useful tools for scanning, interpreting, or predicting macroeconomic activities and/or macropolicy effects in the real world. The general and fundamental assumption is that the economic developments of individual countries and their policy stances can, at least approximately, be measured by small sets of lagging, current and leading indicators. These may be used either as signals to monitor the progress of economies or as a means to strengthen the degree of

European-wide policy cooperation. Such indicators normally can be satisfactorily represented by the realizations of a group of commonly used economic aggregates.

Presumably, there exist a number of generating mechanisms that explain, or fit, the joint observations resulting from scrutiny of these macroeconomic variables and their interrelationships. These should not only be intimately linked with the state of national economies, past, present and future, domestic and abroad, but should also, hopefully, be stable enough over time to warrant further work on trying to master the structure of the world economy.

It is clearly an important task to investigate and come to an understanding of the workings of the underlying mechanisms and to predict and simulate, with the help of economic theories and statistical tools, the likely evolution of the relevant indicators based on such interrelationships. In this regard, the rational expectations hypothesis (REH) advanced by Muth (1961) in the macroeconomic context is often considered optimal [cf. Pain & Britton (1992), pp.82-7, and Wallis (1989), p.44]. This takes into account all available information in the universe in a manner conducive to forming expectations with no systematic errors. Hence, practical Bayesian inference should also be beneficial [e.g., Doan *et al.* (1984), Kadiyala & Karlsson (1993, 1997), Litterman (1980, 1986a), Phillips (1995a), Shoesmith (1990, 1992), West & Harrison (1989), Chapter 4, and Zellner (1985)]. The Bayesian approach can be seen as a step in the same direction, in that it takes into consideration all of our current knowledge, both prior and sample, in a coherent fashion, to facilitate predictions under uncertainty.

However, some words of caution about the use of estimated economic relationships for any empirical prediction and simulation purpose would seem appropriate here. This is especially the case for sudden changes of empirical auto- and/or cross-correlations. Such changes could come about for a wide variety of reasons, for example, those caused by the oil crises in the early 1970s, which led to widespread and significant failures of forecasting models throughout the world. Indeed, on closer examination, the foundations of macroeconomic

systems may be undergoing deep changes caused by artificial stimulation of government pre- or post-election regulation, by deliberate intervention in personal behaviour, and by ordinary socio-economic evolution processes. Such endemic changes of structure will unavoidably constitute major sources of forecast deterioration and hence make the task of forecasting (especially in the long term) much more problematic [Barrell & Pain (1992), p.3].

Fortunately, however, in fundamental and traditional human socio-economic behaviour, there is often more stability than variability. Hence, the deep parameters of the economy with some possible invariants may not change quickly, being highly correlated with current and past episodes and also appearing to exhibit some constant features into the future [see Clements & Hendry (1992a), p.3]. In essence, although the rationality of expectations is still a contentious issue, the logical variant of the famous Lucas critique (Lucas, 1976a), being either testable on specific occasions or manageable in general, does not of itself condemn the forecasting enterprise to failure [see Clements & Hendry (1992a), p.20]. In this and succeeding chapters, therefore, the work attempts to make meaningful inferences and decisions on both theoretical and statistical grounds in the face of newly emerging data and serious intellectual challenges that have emanated from the New Classical macroeconomics.

The main objectives of the research project are clarified in the next section (1.2). The third section (1.3) provides a background, giving a brief overview of some of the aspects and the relative importance of theoretical and econometric developments in macromodel-building strategies.[1] Some relevant key issues surrounding the topic are furnished in the fourth section (1.4). The final section (1.5) contains an outline of the thesis and a flow chart of the research procedure.

[1] A further survey of all the twists and turns in the argument falls outside the scope of this discussion. For a more detailed, highly readable account of the formal unity of a theoretical and a statistical model see Granger (ed) (1991), Chapters 1, 6, 9, 14, 15 & 17.

§1.2 Objectives

The central goal of this thesis is to set up Bayesian vector autoregressive (BVAR) forecasting models for the economies of the United Kingdom, Germany, France and Italy, using IMF, OECD and EC data sources. We investigate whether useful gains, in terms of forecast accuracy, can be made through exploiting the co-movements of some of the principal macroeconomic variables within a European-based multicountry context. The dynamic interactions between the economies specified in the model will be explored, with the aim of shedding some light on complicated questions about whether economic growth paths of different countries converge in some sense over time. For this, a number of separate but related research problems require further analysis.

(a) Is it possible to bridge the gap between the existing theoretical and statistical analyses in the initial specification of an empirical dynamic econometric model, and to proceed in the case where a number of the initial underlying assumptions are rejected?

(b) How can we get round problems of insufficient data and create an adequate data set upon which to base our analyses?

(c) Is it possible to determine long-run constraints, a joint prior structure and its associated prior hyperparameters on the basis of some economic theories, statistical regularities and historical dynamics of the data?

(d) How can we construct, compare and combine models for a given period to generate a general BVAR model subject to certain stochastic and long-run restrictions, and then assess whether the resulting form of the empirical model is an appropriate approximation of the actual data generation process (DGP)?

(e) How can the rationality of economic forecasts be tested in terms of their unbiasedness and efficiency properties?

(f) How can the decision be made about whether the sources of forecast and turning-point errors are cross-country ones or country-specific ones, and how can they be decomposed into avoidable and unavoidable components?

All these substantive issues are of crucial importance for improving our understanding and knowledge of the true nature of the underlying stochastic processes and the evolutionary structure of the contemporary world economy. Through this work, the formulation of more reliable and effective empirical models will be addressed.

§1.3 Background

It has become widely accepted that the role of macroeconomic theory is far more limited than was at first recognized and, whatever the debates, the observed value of data has rarely been challenged. Nonetheless, the impact of rational expectations on the theoretical front has provided an important impetus to the advancement of applied macroeconometrics in general and the development of macroeconometric modelling in particular. In this section, attention will be focused on some basic ideas suggested by theoretical considerations about the long-run equilibrium of the system because, ignoring the basics, any further developments would be rootless.

1.3.1 Macrotheoretical Analysis in Modelling Multivariate Long-Run Equilibrium Interrelationships

In macroeconomics, various pieces of *a priori* theoretical information currently available about long-run, or equilibrium, interlocking relationships among the important aggregates can be attributed variously to Keynesians, monetarists, Bayesians, rational expecters, and so forth. The macroeconomics landscape is littered with economists carrying different banners and offering different opinions. Although, from a methodological viewpoint, the explicit use of *a priori* information is, in principle, highly desirable, the

tasks are how to integrate that information into the inferential process and how to test directly or indirectly the assumptions underlying the economic theories [Freedman (1986), p.127].

On macrotheoretical grounds, the basic conceptual existence of long-run equilibrium relationships, proposed by macroeconomic theory, means that there exists the belief that certain macroeconomic variables should not wander freely or independently of each other; instead, they are expected to move in a specific fashion so that they do not drift too far apart. This is quite consistent with the fact that in the short-run some factors may shock the macroeconomy away from equilibrium, but that this equilibrium will be restored again in the long-run. Although there are many definitions of economic equilibrium, one is described by Machlup (1958, p.9) as follows:[2]

> [an equilibrium is] *a constellation of selected interrelated variables so adjusted to one another that no inherent tendency to change prevails in the model which they constitute.*

The literature on the topic of objective long-run economic hypotheses involving equilibrium concepts is, of course, voluminous and still growing. However, given space limitations, only some general propositions, which I believe have great relevance for econometric practice, will be presented here.

(a) There is a (parabolic) inverse relationship between the going rate of interest and the aggregate demand for money,[3] i.e. lower interest rates will boost money demand in line with the higher money supply. Meanwhile, when interest rates fall, the level of investment in the economy will increase, reflecting government willingness to take action to get the economy growing.

(b) There is a positive linkage between monetary expansion and inflation, meaning that excessive monetary expansion would

[2] See also Chiang (1984), Chapter 3, especially pp.35-6.

[3] For an entertaining account of the approach see Branson (1989), pp.319-46, and Laidler (1985), pp.39-77.

ultimately be channelled entirely into inflation, and eventually the unemployment rate would become immune to whatever is the long-run inflation rate. This is the so-called vertical long-run Phillips curve, in which there are no long-run, but only short-run, changes in (higher) inflation and (lower) unemployment as a result of discretionary monetary intervention [see, *inter alia*, Owen (1986), Chapter 1].

(c) There is a monotonically nondecreasing relationship between total imports and total exports. Thus, the total aggregate income and import levels in the various trading countries would be sufficient to absorb total aggregate exports through merchandise trade flows and prices, and the gains to competitiveness arising from exchange rate depreciation of an economy would also support a rise in its share of world trade. [See, for example, Anderson *et al.* (1992), Artis & Holly (1992), pp.334-5, and Barker *et al.* (1993)].

The Keynesians' prior emphasizes the efficiency of policy and embodies the idea that fiscal policy is of particular significance in the business cycle. This may influence output/(un-)employment through changes in the level of aggregate demand in the standard Hicksian IS/LM model. In contrast, the monetarists' prior stresses the ineffectiveness of policy but takes the view that monetary policy is of central importance in the business cycle and could affect inflation/production through the control of the growth of the (nominal) money supply. This, in turn, can be regarded as a good single index of monetary policy [see Granger (ed) (1991), Chapter 11, and Owen (1986), Chapters 1-3]. As European monetary union gets underway, monetary policy and, in particular, the monetary aggregate indicators will play an important role in the process of transition to full monetary union [see Artis & Lee (ed) (1997), Chapters 1, 2, 12 and 13]. The long-run assertions made by standard macroeconomic theories provide food for thought in the entire modelling exercise.

The traditional role, and much of the recent work of economic theory, is thus formally invoked in four aspects of the initial specification process: the selection of variables; the determination of the signs of unknown parameters; the implications of jointly

simplified dynamic structure; and the imposition of various long-run constraints. Clearly, every economic variable could, potentially, be interrelated or interact with every other variable to some extent. Therefore, apart from special and unlikely circumstances, it is very uncommon to find explanatory variables that appear in equations with exactly zero coefficients if one wants to minimize the predictive expected loss (risk). But when considering a tradeoff between the inclusion of additional variables and the gains of precise description, it is not advisable, in the modelling strategy, that all possible economy-wide variables be included blindly, or tested for inclusion, into the system.

Hence, the first step facing the modeller is to use *a priori* information, without excluding all other information, as a rough guideline for the choice of the observed data series that correspond to the (latent) theoretical variables under scrutiny. The second step is to identify the signs of coefficients, or the direction of change, in a relationship between the variables included [Griffiths *et al.* (1993), pp.796-7]. The third is, at least intuitively, to specify a general weight-declining lag scheme by means of proper priors, so that the influence of past values on current values can be increasingly attenuated. However, see Cooley and LeRoy (1985) for a contrary opinion. Finally, theory is used to influence the estimating form of empirical models, in so far as it evens out the observable short-run fluctuations via its testable long-run restrictions. At an elementary level, whether one accepts the prevalent Keynesian-type and/or monetarist-type prescriptions colours the way in which suitable methods to handle empirical models with these characteristics may be thought of.

The use of complex *a priori* reasoning to guide the specification of a set of desired relationships is in effect only a stepping-stone to a formal dynamic specification. One reason is that economic theories, whether Keynesian or monetarist, are still far from perfectly developed, and should not be expected to give a complete and uniquely specified model. For example, a given economic theory on its own merely tells us that some economic variables are related strongly and others are only weakly related, if at all. It does not

provide the exact form of a prior for the adjustment process and lag structures. Nor is it possible to place sufficient identifying restrictions, long-run or otherwise, on the empirical magnitudes of population parameters of interest so as to derive clear-cut conclusions [see Granger (ed) (1991), Chapter 9].

It is the recognition of the imprecision and incompleteness of macroeconomic theories in practical terms that leads to an emphasis on the important role of quantitative analysis. This recognition also gives rise to the need for blending personal beliefs and historical records into the process of specification search, before a stochastic dynamic model is finally specified.

1.3.2 Macroeconometric Analysis in Modelling Multivariate Short-Run Dynamic Interrelationships

In macroeconometrics, there are essentially two basic views on the complementary roles and relative importance of theoretical structure and data analysis within a dynamic specification framework [see Clements & Mizon (1991), pp.887-8, and Coen *et al.* (1969), pp.152-3]. These contrasting views can have an important effect on the final choice of econometric modelling strategy. At one extreme, a theory-oriented approach, for example, a structural econometric model (SEM), can be regarded as a substantial subject, playing a dominant role in describing the nature of causal links at work or in predicting the effects of a change on the structure of the system. At the other extreme, *a priori* structural hypotheses can be reckoned a humble subject, playing a subsidiary role in any pragmatic and empirical investigation. In this latter view, a data-determined approach, for example, a finite vector autoregressive (VAR) representation that is capable of characterising the frankly *ad hoc* nature of the joint temporal causal structure among a wide range of macroeconomic variables, should have a dominant role [see Sims (1980a)]. On the whole, many researchers hold somewhat intermediate positions by constructing and maintaining theoretically congruent dynamic models.

Serious formulation of macroeconometric models on the world scene

began to take place just after the Second World War, but developed rapidly thereafter. More recently, Sims (1980a, and in later papers) proposed using the heavily parameterized, unrestricted (or loosely restricted) vector autoregressive (UVAR) approach, termed by Cooley and LeRoy (1985) 'atheoretical macroeconometrics', as an important alternative to classical techniques for macroeconometric analysis. In this approach, a vector of jointly endogenous variables is specified to be a linear function of their own and each other's lagged values, not subject to a direct, explicit and restrictive economic theoretic meaning. Since then, this methodology has been widely used for small to medium-sized macroeconometric models in the area of forecasting.

A notable feature of the UVAR model is that the system is a closed one with no contemporaneous variables assigned the extra-model status of exogeneity, no 'incredible' over-identifying restrictions imposed and no trivial endogenous/exogenous distinction allowed. As Anderson (1979) points out, the simple UVAR models should forecast better than structural models since the former, unlike structural models based on more complex and subtle economic theory, do not require complete data and can thus generate forecasts of all variables internally in a fairly straightforward manner. This is especially true in cases where little information is available about the determinants of the vector of variables of primary concern, and a sufficiently large amount of data is obtainable to enable one to produce a linear dynamic 'reduced-form' with a distributed lag of reasonable length.

Although the UVAR models have been extensively utilized, a serious limitation of the approach is the problem of overfitting or overparameterization.[4] Due to multicollinearity and loss of degrees

[4]It has long been recognized that when estimating UVAR models with a relatively large number of variables, without restriction, the number of coefficients, which grows with the square of the number of variables, is large relative to the number of observations typically available, and overfitting occurs. This is likely to result in imprecise estimates of individual coefficients and, hence, many free insignificant coefficients hampering the interpretation of the interrelationships among the variables [see Owen (1986), p.34].

of freedom, overparameterization will generally lead to poor 'out-of-sample' forecasts and less precise estimates of forecast error variance components. One promising forecasting method of tackling overparameterization is the Bayesian vector autoregressive (BVAR) method pioneered, explicitly along Bayesian lines, by Litterman (1980). Instead of adopting hard shape or exclusion restrictions, he imposes the so-called 'Minnesota or Litterman Prior' on the coefficient estimates, which is centred about the vector, or multivariate, random walk process. The idea met with considerable scepticism, until Litterman's model was found to outperform other forecasts of many economic variables [Doan (1996), p.8-17, Granger (ed) (1991), p.3, and Shoesmith (1992), p.93].

As Todd (1984) clarifies, a basic BVAR model might seem to resemble a UVAR model in the formation of its equations and also resemble a structural model in its use of priors to prevent overfitting. However, from a Bayesian perspective, both the UVAR and SEM approaches are either too vague (in the former case) or too extreme (in the latter case), and thus probably either overstate or understate the modeller's true beliefs. The BVAR modeller holds that more recent lags of a vector of variables are more relevant in forecasting than now distant ones, whereas the UVAR constructor pretends to be completely ignorant and treats all possible lag structures to be data determined. At the same time, the BVAR modeller uses prior beliefs to specify the 'guesstimated' coefficients, while allowing the data to override each of the guesses in the light of sample evidence afterwards. This method is in contrast to the approach of the SEM builder, who pretends either to be absolutely confident about the zero excluded coefficients, no matter what the historical data suggest, and to be absolutely ignorant about the posited included coefficients, no matter what the modeller actually believes.

A distinctive feature of the prototypical BVAR method is that it permits forecasters to incorporate, in a logically consistent manner, both prior and sample information through Theil's mixed estimation technique, with the best settings of the hyperparameters to generate international forecasts or to suggest macro policies.

This is important because it gives investigators a flexible, credible way of expressing personal beliefs more accurately and a standard, objective procedure for combining those beliefs with historical record more formally.

The BVAR prototype is inherently a naive backward-looking model, since the default of the Minnesota prior, which effectively shrinks all the parameters either exactly or approximately to the vector random walk (with or without drift) hypothesis, is based on the implicit 'business as usual' assumption.[5] In other words, it is assumed that the behaviour of many macroeconomic variables will be the same in the future as it was in the sample period, so that, for such a variable, the best forecast of its future values will simply equal its current value. This is, perhaps, a chief source of challenge for the model-based forecasters, since these 'no-change' forecasts can be extremely difficult to improve upon.

In the search for a more generic, economically meaningful model, the BVAR prototype has been readily extended to an open system for the dynamic, disequilibrium adjustment process based on the error-correction mechanism (ECM). The emphasis is on the use of a further exogenous factor, the equilibrium error, which arises primarily from the concept of multi-cointegration. According to the relationship between cointegration and error correction expounded by Granger and his co-authors (see e.g. Granger, 1986b, for references), the resulting BVAR model, very much like a negative feedback network, is essentially a self-correcting model. The impact of a long-run or equilibrium solution, as implied by *a priori* theory, can be introduced into the model in order to restrict the short-run dynamics, and the forecast can be set back on track, so that under- or over-prediction of an equation in the past will automatically lead to a commensurate adjustment in the future. This relatively inexpensive BVAR model, with both prior and cointegration constraints, will be presented as an effective means of generating accurate forecasts and providing a tough benchmark of comparison for forecasts derived in more traditional ways.

[5] For details of this literature, see Artis (1988), pp.25-6.

In order to apply state-of-the-art time series econometrics modelling techniques that incorporate prior information to the major macroeconomic indicators to generate forecasts, some relevant key issues surrounding the topic should first be considered in a critical light.

1.4.1 Bayesian Analysis and Retrospection

An overview of the evolution of the macroeconometric approach has indicated that there is a long and unfortunate tradition of placing too much emphasis on macroeconomic theory, as opposed to historical evidence, in the specification and evaluation of macroeconometric models. The spirit of this tradition was to a large extent rooted in the desire to search for a true underlying model by emulating the approach of the harder sciences [see Bodkin *et al.* (1991), Chapters 6-8, and Granger (ed) (1991), Chapters 6 and 15].

However, in many situations it is perhaps naive to argue that an economic theory or model is 'true' or 'false' *per se*; it may be more reasonable to argue that a theory or a model is 'useful' or 'not useful' in facilitating the specification of empirical relationships, or in illuminating the particular phenomena of interest. The contention here is that the crucial assumptions on which the conclusions of a theory depend sensitively are never precisely true, the links between theories and empirical specifications in macroeconometrics are not very close, and the available observations are not sufficient to enable one to estimate all the 'true' parameters with great accuracy. It is, therefore, counterproductive to think that one could finally achieve some sort of complete, once-and-for-all, 'true' mechanism in the longer term.

Recent developments in macroeconometrics have changed interests towards the use of both data and theory, sophisticatedly simple hypotheses or methods, prediction and model selection criteria and the Bayesian learning procedure in model specification and

evaluation. It seems to be generally agreed that in building econometric models, theory should not be binding on the model, but should be true within the model [Granger (ed) (1991), pp.18-9]. The 'bottom line' is how well models perform in interpreting what is going on, or in predicting what is going to happen. In this respect, the explicitly Bayesian shrinkage techniques have had a significant impact on econometric studies, providing a formal framework for handling background or prior information under uncertainty, and will undoubtedly open up a new vista for the progress of econometric modelling [Todd (1984)].

1.4.2 Bayesian Modelling and Estimation

Linearity, as a useful approximation to many non-linear problems, has long been used in macroeconometrics to model the dynamic interrelationships between principal macroeconomic variables.[6] We will first survey the previous forecasting record of the UK, Germany, France and Italy. Attention then turns to the construction of macroeconometric models that incorporate both prior and cointegrating restrictions for these countries, within a similar multivariate linear, or log-linear, stochastic framework. Through the foreign trade connections between the economies, these various country models can then ultimately be combined into a general BVAR model for the European economies.

Along the way, a number of relevant problems and issues will be investigated, such as the inspection of cointegrating vectors; the refinement of lag structures; the exploration of causal linkages; the treatment of deterministic terms; the estimation of the sensitivity of forecast performance to prior information; the employment of error-correction mechanisms; and the determination of the best settings of the prior hyperparameters. The extensive data set used is obtained from various IMF, OECD and ONS databanks. Rolling mixed-estimations with the aid of the Kalman Filter updating procedure will be used for the BVAR models, thereby allowing

[6]For references to the forecasting literature for non-linear models see, among others, Mariano (1985) and Mariano & Brown (1983, 1991).

parameters to vary over time. The resulting t, F and R^2 values will be presented and the comparative forecast results in terms of Theil's U statistics (or RMSEs) will be reported. All estimation is conducted using the Microfit, MicroTSP, PcGive and RATS software packages on an IBM (or compatible) personal computer (PC).

1.4.3 Bayesian Forecasting and Evaluation

The prediction of the future time paths of macroeconomic variables is one of the most noticeable manifestations of the use of macroeconometric models [Stewart (1986), p.264]. In this study, both single- and multi-country forecasts will be made and evaluated from 1991Q1 onward for a similar range of variables. This provides a rigorous testing ground for the BVAR models and their uses.

1.4.3a *Multicountry Forecasts*

Frequently, the term forecast is used to denote a statement about future events, whereas prediction is used to define an implication of a model. Ex ante forecasting with an econometric model involves the use of past and current information available prior to the forecast period in order to generate a joint predictive distribution, together with a final forecast for future values of the series that have, as yet, not occurred. In this exercise, both simulated and genuine 'out-of-sample' quarterly forecasts will be produced mainly for the short-term (one to eight quarters ahead of the date of occurrence) and the appropriate BVAR forecasting models will be selected to form optimal (linear) forecasts, subject to as little error as possible. In addition, the multicountry forecasts generated by the BVAR models are unconditional, since they do not depend on explicit assumptions about the future course of the projected external determinants of economic activity [McNees (1986), p.15].

1.4.3b *Forecast Evaluation*

A statistical evaluation of the quality of forecast performance of the BVAR models through a comparison with the UVAR and other

conventional models over the forecast period will be presented by using Wald parametric encompassing tests in truly ex ante circumstances. This permits assessment of whether the BVAR approach derived with proper prior and cointegration constraints is superior, on average, to other traditional techniques. The encompassing of econometric hypotheses enables us to go beyond simple Root Mean Squared Error (RMSE) comparisons of forecast evaluation so as to achieve a more generally applicable and effective model. Meanwhile, the forecast and turning-point errors from actual outcomes will be analysed to evaluate the efficiency and bias of individual BVAR forecasts in a changing economic environment.

1.4.3c *Policy Simulations*

Probably the most formidable critique to the quantitative policy evaluation proposition would appear to originate with the Muthian rational expectations hypothesis and the related Lucas (1976a) critique of mainstream macroeconometric model-building. The internal logic of REH is that sensible economic agents will use all the available information efficiently in order to avoid systematic errors in their expectations, or forecasts. The implication is that, given an instantaneous market clearing paradigm, rational agents act to eliminate any effects of systematic policy on real factors. Hence, only unexpected monetary or fiscal policy changes can have any real effects. This is the neutrality or policy ineffectiveness proposition. In practice, a shock can provoke enough variation in aspects of the world economy to allow modelling via data on multiple time-series. If particular economic variables do not vary, then it is impossible to measure their impact on economic activity using only time-series information.

The rational-expectations logic of the Lucas critique is that structural behavioural equations will not be invariant to alternative policy rules and regimes. This implies that rational individuals ought to understand the nature of the policy in force and use this information in forming expectations. Hence, were a new set of policy proposals to be implemented, the parameters of the system would shift systematically. Thus, the structure would break

down and, in particular, would no longer permit a sound prediction of the likely consequences of a new policy action. This is known as Goodhart's (1986) law in the monetary area. It would have represented a forceful criticism of macroeconometric modelling had the argument been sustained.

However, the quantitative significance of the Lucas critique for policy analysis is an open question. One can strongly refute the argument in so far that the likely size of the error from such a criticism, for many experiments, may not be very great.[7] If this were the case, then there would be a return of confidence in the use of simulation techniques for minor policy changes. In reality, there is nothing very special about rational expectations theory that proves that REH is the only way to interpret the behaviour of agents. The developments of the New Classical macroeconomics require a strategy of explaining economic activity in terms of rational individual choice. If there were problems in which there were no rational choices, then this strategy would not equip the New Classical approach to analyse them at all, thereby limiting its domain. Even in the absence of any such limitations, the New Classical analysis would still be unable to offer a complete explanation of economic behaviour. Accordingly, while the challenge raised by the New Classical school has posed some important questions about the dynamics of the macroeconomy, it is reasonable to doubt whether this critique is applicable in its present form.

In this work, we will stick to the paradigm that behavioural parameters will not suffer greatly from the Lucas criticism, and use the formulated model to learn about the dynamic interactions between (among) two (or more) of Europe's major economies.

§1.5 Intended Scope and Flow Diagram

The format of this thesis is, therefore, made up as follows. After

[7] See, for example, Artis & Holly (1992), pp.336-7, Bodkin *et al.* (eds) (1991), pp.551-5, Granger (ed) (1991), Chapters 8-11, and Wallis (1989), pp.35-8.

the brief discussion of recent theoretical and practical developments in depicting the interrelationships between a number of macroeconomic variables, particular attention is paid, in Chapters 2 and 3, to the prototypical BVAR modelling. This model is constructed in the framework of standard stationary normal linear dynamic systems, without much loss of generality and with a noticeable benefit in simplicity and clarity.

The set-up of a more general BVAR model subject to both prior and cointegration restrictions is then discussed in Chapters 4 and 5 for non-stationary time series. A pragmatic, systematic BVAR and/or BVAR-EC forecasting model selection procedure, which can readily be implemented on available econometric packages, is put forward.

In Chapters 6 and 7, the transformation of the macro data for the formulation, modification and selection of empirical BVAR forecasting models is laid out. Various ex post and ex ante quarterly forecasts based on the mix-estimated BVAR and BVAR-EC models are appended and displayed in the accompanying graphs over the forecast period, 1991Q1-1998Q4, for Europe's leading industrial economies. The relative efficiency and accuracy of BVAR and BVAR-EC models are evaluated in detail using an appropriate forecasting criterion (e.g. Theil's U statistic). Typically, the different sorts of forecast or turning point errors are analysed, in order to assess the real causes of failure in a multinational forecasting context.

To provide an overview, the outline of the proposed methodology is schematically represented in Fig.1.1. The whole system depicted in the figure can be perceived as the exploitation and treatment of plausible information flows, as well as the unification and refinement of theoretical and empirical analyses.

Summaries, practical conclusions and some further thoughts and priorities for future work are set forth in Chapter 8. In addition, the tabulation of experimental results, specifications of computer programs stored in ASCII files ending with suffix *.ASC and other details are provided in Appendices A-F, followed by references.

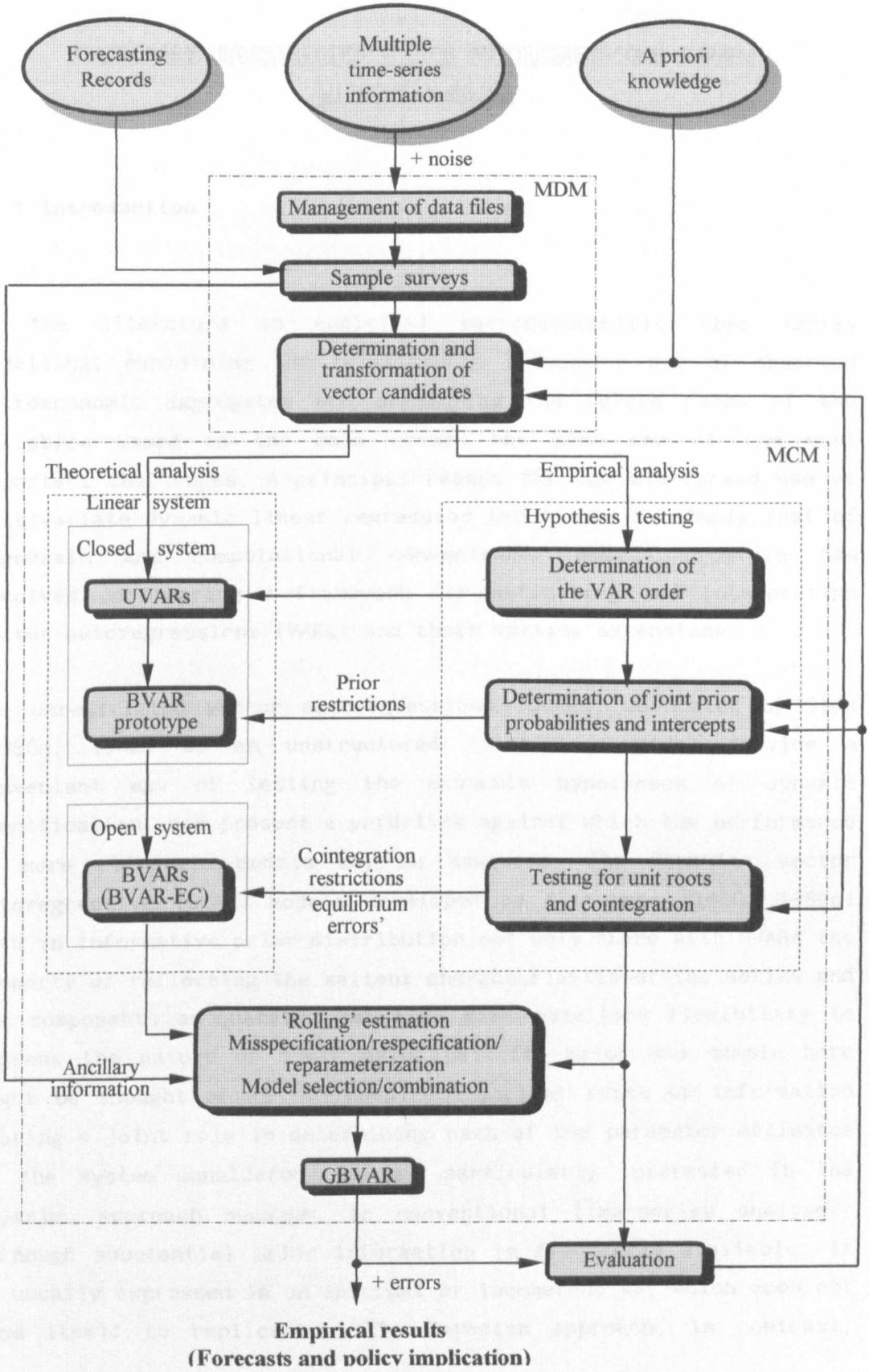

Fig.1.1: The methodological framework of the improvement and selection of a general BVAR model

CHAPTER 2

STATIONARY UNRESTRICTED VECTOR AUTOREGRESSIONS (UVARs) WITHOUT A PRIOR

§2.1 Introduction

In the literature on empirical macroeconometric time series modelling, explaining the relationship between a set of observed macroeconomic aggregates and predicting the future paths of the variables based on the data chosen has been one of the most important challenges. A principal reason for the widespread use of multivariate dynamic linear regression techniques is simply that of algebraic and computational convenience. Recent progress has involved considering a framework for estimating and interpreting vector autoregressions (VARs) and their various extensions.

The unrestricted vector autoregressions (UVARs) advocated by Sims (1980a, 1982) as an unstructured first-stage model provide a convenient way of testing the economic hypotheses of dynamic specification, and present a yardstick against which the performance of more elaborate models may be assessed. The Bayesian vector autoregressive (BVAR) models developed by Litterman (1980, 1986a) with an informative prior distribution not only share with UVARs the property of reflecting the salient characteristics of the series and its components adequately, but also give modellers flexibility to express the nature of their prior beliefs. Prior and sample here might be thought of as two equally important types of information playing a joint role in determining each of the parameter estimates in the system considered. We are particularly interested in the Bayesian approach because, in conventional time-series analyses, although substantial prior information is frequently available, it is usually expressed in an informal or incoherent way which does not lend itself to replication. The Bayesian approach, in contrast,

provides a general formal method for handling proper prior and finite sample information in a coherent methodological framework of macroeconometric modelling [Zellner (1985), p.255].

This chapter starts by reviewing some of the important concepts involved. It then goes on to introduce the approach employed in the construction of linear time domain models. The chapter concludes by illustrating how to construct a basic, finite-order, stationary BVAR forecasting model with stochastic prior restrictions.

§2.2 A Linear Stochastic Multivariate System for Nonintegrated Time Series

From a theoretical viewpoint, the multiple time-series model could always be embodied within the appropriate structure of a multivariate econometric model. A properly specified multivariate econometric model should provide a richer pattern of correlations and yield more accurate forecasts than a purely time-series model. The main objective of empirical econometric modelling is to find a relatively simple specification which, in broad terms, satisfactorily captures the characteristics of the observed data. To that end, the work in this study will concentrate initially on a jointly stationary VAR representation (possibly after differencing) of a set of economic data. For expository purposes, we will take the simple stationary case first.

2.2.1 Stationary Vector Autoregressive (VAR) Processes and Their Properties

In conventional (multiple) time-series macroeconometric analyses, a natural starting-point is to treat a discrete vector stochastic process as the actual underlying mechanism which generates observations on a given set of random variables, say $\underline{Y}_t=(Y_{1t},\ldots,Y_{Kt})'$, $t=1,\ldots,T$. Each observation is one realization of the vector stochastic process, also denoted by $\{\underline{Y}_t\}$ when there is no possibility of confusion.

In principle, the vector stochastic process may be characterised by the joint distribution function of all finite, time-ordered subsequences of random vectors $\{\underline{Y}_t,\ t\in T_n \subset N\} \longrightarrow \mathbb{R}^K$, K-dimensional Euclidean space. Thus, if we could somehow specify the joint probability distribution function $p(\underline{Y}_1,\ldots,\underline{Y}_T)$ for our process, we would then fully capture the true random behaviour associated with future outcomes. In practice, however, empirical analyses are rarely based on the complete distributions, but on parameters of the distributions, such as the mean vectors $E[\underline{Y}_t]$ and covariances $E[(\underline{Y}_t - E[\underline{Y}_t])(\underline{Y}_{t-s} - E[\underline{Y}_{t-s}])']$. However, to be able to infer all the values of these first and second moments from just a single realization, some further assumptions have to be made.

One important assumption generally made is that of ergodicity. It asserts that the random sample moments approach their fixed population counterparts as the available sample size T goes to infinity.[1] A second assumption is that of stationarity, which has, in general, two conceptually distinct but closely related definitions: strict-sense stationarity; and wide-sense stationarity.

Strict stationarity implies that both the joint multivariate density function $\{p(\underline{Y}_t,\ldots,\underline{Y}_{t+h}),\ h\neq 0\}$ and the marginal function (h=0) are independent of a displacement in time origin and that any set of h+1 observations has identical distributions. Wide-sense or covariance stationarity implies that the first two moments of the joint distributions are finite and time invariant; that is,

$$E[\underline{Y}_t] = \mu_Y < \infty, \qquad \forall t,$$

and

$$E[(\underline{Y}_t - \mu_Y)(\underline{Y}_{t-s} - \mu_Y)'] = \begin{cases} \Sigma_Y = \Gamma_Y(0) < \infty, & \forall t,\ s = 0; \\ \Gamma_Y(s) < \infty, & \forall t\ \&\ s,\ s \neq 0. \end{cases}$$

Thus all such moments exist and that they do not depend on the particular time point t but only on the absolute value of the time

[1] See, *inter alia*, Hannan (1970), pp.200-20, for a more thorough discussion of ergodicity.

difference, s. The practical consequence of this property is that it allows us to study the sampling properties of the variables of interest through summary measures like sample moments based on a single realization.[2]

For a jointly normally distributed vector stochastic process, often known as a Gaussian vector process (determined entirely by its first two moments), weak and strict stationarities are equivalent [Harvey (1992), pp.49-50]. For example, a vector white-noise process $\{\underline{u}_t=(u_{1t},\ldots,u_{Kt})',\ t\in T\}$, i.e. one with $E[\underline{u}_t]=0$, $E[\underline{u}_t\underline{u}_t']=\Sigma_u$(nonsingular)$<\infty$ and $E[\underline{u}_t\underline{u}_s']=0$ for $s\neq t$, is clearly stationary. This process may be normally independently distributed (NID) or, less restrictively, independently identically distributed (IID) with zero mean vector and finite covariance matrix. In the former case strict and wide sense stationarities hold, whilst in the latter wide sense stationarity certainly holds.

A standard time series modelling strategy is to assume the current value of each variable Y_{jt} (assumed stationary for the moment), $j\in[1,K]$, to be affected by past values of itself and the past (and possibly also the current) values of other variables that are present, together with a random disturbance at each point in time. That is,

$$\begin{aligned}
Y_{1t} &= D_1 - b_{12,0}Y_{2t} - \ldots - b_{1K,0}Y_{Kt} + b_{11,1}Y_{1,t-1} + \ldots + b_{1K,1}Y_{K,t-1} \\
&\quad + \ldots + b_{11,p}Y_{1,t-p} + \ldots + b_{1K,p}Y_{K,t-p} + \epsilon_{1t}; \\
&\quad \ldots\ldots \\
Y_{Kt} &= D_K - b_{K1,0}Y_{1t} - \ldots - b_{KK-1,0}Y_{K-1,t} + b_{K1,1}Y_{1,t-1} + \ldots + b_{KK,1}Y_{K,t-1} \\
&\quad + \ldots + b_{K1,p}Y_{1,t-p} + \ldots + b_{KK,p}Y_{K,t-p} + \epsilon_{Kt}, \qquad (2.1)
\end{aligned}$$

where the terms ϵ_{jt}, $j=1,\ldots K$, are assumed serially uncorrelated white-noise disturbances with constant standard deviations σ_1 through σ_K, respectively.

[2] For details see Judge *et al.* (1985), Chapters 7 and 16, and Pindyck & Rubinfeld (1991), Chapters 15 and 16.

These K equations constitute a (primitive) K-dimensional vector autoregressive process of order p or VAR(p). The structure of the system incorporates feedback since current values Y_{1t} to Y_{Kt} are allowed to affect each other. For example, $-b_{1K,0}$ is the effect of a unit change in Y_{Kt} on Y_{1t}. If $b_{1K,0}\neq 0$, then Y_{Kt} has a contemporaneous effect on Y_{1t}, and ϵ_{Kt}, being pure shocks in Y_{Kt}, has an indirect contemporaneous effect on Y_{1t} [see e.g. Enders (1995), pp.294-7].

Due to the feedback effects in the system, (2.1) cannot be estimated directly using OLS, since the regressors are correlated with the error term. However, it is possible to transform the system into a reduced form. We can write the system in the compact form as:

$$\begin{aligned}\underline{B}_0\underline{Y}_t &= \underline{D} + \underline{B}_1\underline{Y}_{t-1} + \ldots + \underline{B}_p\underline{Y}_{t-p} + \underline{\epsilon}_t \\ &= \underline{D} + \sum_{i=1}^{p}\underline{B}_i\underline{Y}_{t-i} + \underline{\epsilon}_t; \quad \forall t, \end{aligned} \tag{2.2}$$

where

$$\underline{Y}_t = \underset{(K\times 1)}{\begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{Kt} \end{pmatrix}}, \quad \underline{D} = \underset{(K\times 1)}{\begin{pmatrix} D_1 \\ \vdots \\ D_K \end{pmatrix}}, \quad \underline{B}_0 = \underset{(K\times K)}{\begin{pmatrix} 1 & b_{12,0} & \cdots & b_{1K,0} \\ \vdots & & \ddots & \vdots \\ b_{K1,0} & \cdots & b_{KK-1,0} & 1 \end{pmatrix}},$$

$$\underline{B}_i = \underset{(K\times K)}{\begin{pmatrix} b_{11,i} & \cdots & b_{1K,i} \\ \vdots & \ddots & \vdots \\ b_{K1,i} & \cdots & b_{KK,i} \end{pmatrix}}, \; i = 1, \ldots, p, \text{ and } \underline{\epsilon}_t = \underset{(K\times 1)}{\begin{pmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Kt} \end{pmatrix}}.$$

Premultiplication by $\underline{B}_0^{-1}$ allows us to obtain the equivalent vector autoregressive (VAR) model in standard form. The current observation of $\underline{Y}_t$ is a linear aggregation (or linear filter) of its own p past values plus constant and random error terms:

$$\begin{aligned}\underline{Y}_t &= \underline{C} + \underline{A}_1\underline{Y}_{t-1} + \ldots + \underline{A}_p\underline{Y}_{t-p} + \underline{u}_t \\ &= \underline{C} + \sum_{i=1}^{p}\underline{A}_i\underline{Y}_{t-i} + \underline{u}_t; \quad \forall t. \end{aligned} \tag{2.3}$$

Here $\underline{A}_i=\underline{B}_0^{-1}\underline{B}_i$, $i=1,\ldots,p$, are $K\times K$ matrices of lag coefficients,

$\underline{C}=\underline{B}_0^{-1}\underline{D}$ is a K×1 vector of deterministic terms relating to the mean of the stochastic process, and $\underline{u}_t=\underline{B}_0^{-1}\underline{\epsilon}_t$ is assumed a K×1 Gaussian vector white-noise process, denoted $\underline{u}_t \sim NID(0, \Sigma_u)$.[3]

To distinguish between the systems represented by (2.2) and (2.3), the first is called a structural VAR (or a VAR in primitive form) and the second a VAR in standard form.[4] Only the standard form VAR model can be estimated using OLS.[5] To ensure that the same linear generation law prevails even outside the sample period, we assume that the model extends backwards and forwards in time, and use the expression '$\forall t$' in the generating equations of the $\underline{A}$ weights.

It is useful to introduce the lag, or backshift, operator L such that $L^n\underline{Y}_t \equiv \underline{Y}_{t-n}$ (and $L^n\underline{C}\equiv\underline{C}$). Using this operator, Eq.(2.3) can be rewritten as

$$\underline{A}_p(L)\underline{Y}_t = \underline{C} + \underline{u}_t, \quad \underline{u}_t \sim NID(0, \Sigma_u), \quad \forall t, \tag{2.4}$$

where the lag polynomial of order p

---

[3]The variance-covariance matrix (Σ_u) can be written as:

$$\Sigma_u = var(\underline{u}_t) = E(\underline{u}_t\underline{u}_t') = E[(\underline{B}_0^{-1}\underline{\epsilon}_t)(\underline{B}_0^{-1}\underline{\epsilon}_t)'] = (\underline{B}_0^{-1})E(\underline{\epsilon}_t\underline{\epsilon}_t')(\underline{B}_0^{-1})'$$
$$= (\underline{B}_0^{-1})\,diag(\sigma_1^2,\ldots,\sigma_K^2)\,(\underline{B}_0^{-1})',$$

which is, in general, a nonsingular K×K matrix (with time-invariant variance and covariance terms). Hence, we get

$$\Sigma_u = \begin{pmatrix} var(u_{1t}) & \cdots & cov(u_{1t},u_{Kt}) \\ \vdots & \ddots & \vdots \\ cov(u_{1t},u_{Kt}) & \cdots & var(u_{Kt}) \end{pmatrix} \equiv \begin{pmatrix} \sigma_{11}^2 & \cdots & \sigma_{1K} \\ \vdots & \ddots & \vdots \\ \sigma_{K1} & \cdots & \sigma_{KK}^2 \end{pmatrix},$$

where $\sigma_{jj}^2 = var(u_{jt})$ and $\sigma_{ji}=\sigma_{ij}=cov(u_{jt},u_{it})$; $i,j\in[1,K]$.

[4]For simplicity, when we say a 'VAR' process we mean a 'standard VAR' in the following analysis.

[5]However, it is essential to realize that, normally, it is impossible to recover all of the information present in the primitive system from the OLS estimates of (2.3), unless we are willing to restrict appropriately the primitive system using block exogeneity restrictions. We return to these issues in the next section. [For further details, see Enders (1995), pp.294-305.]

$$\underline{A}_p(L) = I_K - \underline{A}_1 L - \ldots - \underline{A}_p L^P.$$

The focus of interest is often the parameter matrices $\underline{A}_i$, i=1,...,p, not $\underline{C}$ which is constant for each individual. Assuming $\underline{C}$=0 with no loss in generality,[6] we have

$$\underline{A}_p(L)\underline{Y}_t = \underline{u}_t$$

or

$$\underline{Y}_t = \underline{A}^{-1}_p(L)\underline{u}_t.$$

It follows that the stationary linear autoregressive process may be thought of as the output $\underline{Y}_t$ from some linear backward-looking filter with transfer function given by the inverse of the autoregressive operator $\underline{A}_p(L)$, when the input is vector white noise $\underline{u}_t$. Also, the transfer function of the linear filter relating $\underline{Y}_t$ to $\underline{u}_t$ can be viewed as a 'black box', whereby the nature of the input-output relationships is estimated with no explicit support from theory.

2.2.1a *Stationarity in VAR (p) Processes*

Stationarity in VAR(p) processes is summarized in the following proposition without proof.[7]

PROPOSITION: A necessary and sufficient condition for $\underline{Y}_t$~VAR(p) process to be stationary is that

$$\det[\underline{A}_p(\xi)] = |\underline{A}_p(\xi)|$$
$$= |I_K - \underline{A}_1\xi - \ldots - \underline{A}_p\xi^p| \neq 0, \quad |\xi| \leq 1. \qquad (2.5)$$

This condition provides a convenient means for checking stationarity on a given set of observations in principle or testing a null

[6]Taking $\underline{C}$=0 is equivalent to scaling $\underline{Y}_t$ in terms of deviations about its (constant) mean, given that the process is stationary. For further details, see Appendix B.

[7]For a proof of the proposition see, among others, Enders (1995), pp.412-8.

hypothesis on the boundary between stationarity and nonstationarity in practice. There are three possible cases to note about the proposition.

Case (a): Suppose $|\underline{A}_p(\xi)| \neq 0$, for $|\xi| \leq 1$, a stationary case with condition (2.5) fulfilled. In this case the weights given to past shocks will diminish with time; that is, the present is more important than the past.

Case (b): Suppose $|\underline{A}_p(\xi)| = 0$, for $|\xi| = 1$, a unit root case with one or more unit roots and all other roots outside the complex unit circle. This is a specific class of nonstationary case with important economic and statistical properties [see Fuller (1976), Chapter 8, especially pp.366-82]. In this case the weights given to distant shocks will be as important as those given to more recent ones, implying that the past and present are equally important.

Case (c): Suppose $|\underline{A}_p(\xi)| = 0$, for $|\xi| < 1$, an explosive case with at least one of the roots (strictly) inside the unit circle. In this case the time-changing mean and variance of the process would drift further and further away, at an exponential rate, from any fixed reference point [Pindyck & Rubinfeld (1991), p.478], and weights given to previous shocks will be greater than those given to current ones. In other words, the past is more important than the present. This is not often characteristic of typical economic phenomena.

Thus, only the first two cases would seem to be of much practical interest in economic analysis. Case (b) can often be transformed into a stationary one, either by taking differences or by imposing special restrictions called cointegration restrictions. These processes will be discussed in detail in Chapters 4-6.

2.2.1b *Wold's Decomposition Theorem*

Wold's decomposition theorem (Wold, 1938) states that any stationary multivariate stochastic process $(\underline{Y}_t)$ can be uniquely and additively decomposed into two mutually orthogonal component subprocesses: a

linearly deterministic (or nonstochastic) part $(\underline{Y}_t^*)$, and a purely nondeterministic (or stochastic) part $(\underline{Y}_t^\dagger)$.[8] By deterministic, we mean that the subprocess can be predicted perfectly from its own past alone, at any observation point. By nondeterministic, we mean that the subprocess, containing no deterministic components, can be represented as a linear combination of a sequence of uncorrelated or orthogonalized innovations in the equations of the system. That is,

$$\underline{Y}_t = \underline{Y}_t^* + \underline{Y}_t^\dagger = \underline{D}_t + \sum_{i=0}^{\infty} \underline{\Phi}_i \underline{u}_{t-i} = \underline{D}_t + \underline{\Phi}(L)\underline{u}_t, \qquad (2.6)$$

where $\underline{D}_t$ is a K×1 vector of deterministic components, the $\underline{\Phi}_i$'s are K×K matrices of time-invariant parameters, $\underline{\Phi}_0$ is an identity matrix and the matrix polynomial, $\underline{\Phi}(L)$, is absolutely summable. The K×1 error vector $\underline{u}_t$ is, as usual, a multivariate white-noise process.

Since it is assumed throughout that in a covariance stationary process all parameters are time invariant, the deterministic term $\underline{D}_t=\mu_Y$, bearing in mind that $\underline{D}_t\neq\mu_Y$ will imply time-dependent levels or polynomial trends, and will violate one of the conditions for second-order stationarity. We then neglect $\underline{D}_t=\mu_Y$ in the following analysis for ease of notation (though when dealing with real data, its presence can be important). At the same time, the matrix $\underline{\Phi}(L)$ in the purely nondeterministic term can often be adequately approximated by the product of two matrices, $\underline{A}_p^{-1}(L)$ and $\underline{M}_q(L)$, each involving only parsimonious lag polynomials in the sense that p and q are relatively small.

One is then led to consider probably the most important family of linear schemes in time-series modelling

$$\underline{A}_p(L)\underline{Y}_t = \underline{M}_q(L)\underline{u}_t$$

or

$$\underline{Y}_t = \sum_{i=1}^{p} \underline{A}_i \underline{Y}_{t-i} + \underline{u}_t + \sum_{i=1}^{q} \underline{M}_i \underline{u}_{t-i},$$

[8] A formal discussion of this theorem may be found in, for example, Brockwell & Davis (1991), Chapter 5, Granger & Newbold (1986), Chapters 1, 2 and 7, Harvey (1992), Chapters 2 and 8, and Lütkepohl (1991a), Chapter 2.

where the respective vector AR and MA operators are

$$\underline{A}_p(L) = I_k - \underline{A}_1 L - \ldots - \underline{A}_p L^p,$$

and

$$\underline{M}_q(L) = I_k + \underline{M}_1 L + \ldots + \underline{M}_q L^q.$$

This process is called a K-dimensional vector mixed autoregressive-moving average process of order (p, q) or VARMA(p, q). To guarantee both stationarity and invertibility of this process we require that

$$|\underline{A}_p(\xi)| = |I_K - \underline{A}_1\xi - \ldots - \underline{A}_p\xi^p| \neq 0, \qquad |\xi| \leq 1,$$

and

$$|\underline{M}_q(\xi)| = |I_K + \underline{M}_1\xi + \ldots + \underline{M}_q\xi^q| \neq 0, \qquad |\xi| \leq 1.$$

With these two separate conditions fulfilled there is a unique VARMA representation of a vector stochastic process corresponding to the autocorrelation structure.[9] The estimation of a VARMA process inevitably requires a nonlinear procedure owing chiefly to the presence of the MA components, which complicate analysis somewhat.

Wold's decomposition theorem implies that every purely nondeterministic, wide-sense stationary process can be written as an infinite MA representation which, in general, can be inverted to form a VAR representation of infinite order [see Lütkepohl (1991a), p.20]. The stationarity of an infinite VAR, however, requires that the VAR weights die out gradually [see Brockwell & Davis (1991), especially 5.7, for a further discussion]. Hence, it is usually the case that the true generating mechanism of the process under consideration can be adequately approximated by a finite order VAR process with lag length truncated at some suitable p

$$\underline{Y}_t = \sum_{i=1}^{p} \underline{A}_i \underline{Y}_{t-i} + \underline{u}_t, \tag{2.7}$$

or

$$\underline{A}_p(L)\underline{Y}_t = \underline{u}_t; \quad \text{with } E[\underline{u}_t\underline{u}_t'] = \Sigma_u. \tag{2.8}$$

[9] For a fuller discussion of the formalized conditions of the uniqueness of ARMA representations along with various equivalent formulations, see Judge *et al.* (1985), pp.658-9.

This process constitutes a valid basis for representing a broad range of stationary processes within a single class of parsimonious multivariate autoregressive processes [also see Lütkepohl (1991a), p.20]. As such, it is a powerful result and represents an efficacious simplification in applied work.[10]

In practice, the true DGP will not be known or be constant through time and the hypothesized constant-parameter model will not coincide with that DGP. However, for developing technical analyses, these quite restrictive assumptions are often required to ensure tractability.[11] This leads naturally on to the question of identification of a suitable parsimonious representation of the actual generating mechanism, guided by both *a priori* theory and empirical observation.

2.2.2 Concepts of Exogeneity and Causality

Usually, in the study of relationships between time series, it is useful to invoke the concepts of exogeneity and Granger-causality to make the analysis meaningful. These concepts, as characterized by Engle *et al.* (1983), and Engle and Hendry (1993), serve different purposes. Exogeneity is defined in terms of a group of variables for the purpose of conducting inference about parameters of interest, while causality is introduced by reference to a series' importance in prediction for the purpose of producing an operational testing procedure.

---

[10] In theory, any stationary K-dimensional VAR(p) process can be further stacked into a corresponding stationary Kp-dimensional VAR(1) form after a suitable change of notation. Also, as a direct consequence of weak stationarity, the first two moments (means, autocovariances and autocorrelations) will, amongst other things, exist and remain constant. For their derivations in a stationary VAR(p) process see Lütkephol (1991a), Chapter 2.

[11] See, for example, Clements & Hendry (1992a), p.1, and Judge *et al.* (1985), p.660.

2.2.2a *Weak and Strong Exogeneity*

The classical dichotomy between endogenous variables (observable outputs explained inside the system) and exogenous variables (observable inputs determined outside the system), in such a way that they are orthogonal to random disturbances (unobservable inputs to the system under investigation), highlights the properties of variables in a control context. It also provokes many interesting and important problems in practical macroeconometrics. But the decision as to which variables are to be treated as exogenous depends on economic reasoning, not on the mathematical statement of any conditional model. Essentially, the validity of simplifying conditional forecasts from the joint distributions of observable variables is explicitly based on the legitimacy of classifying the variables into weakly and strongly exogenous groups [see Engle *et al.* (1983), and Engle & Hendry (1993)].

To formalize, note that the joint distribution of the observed variables $\underline{Y}_t=(\underline{X}_t', \underline{Z}_t')'$, conditional on the past, can always be factorized as the conditional distribution of $\underline{X}_t$ given $\underline{Z}_t$ times the marginal distribution of $\underline{Z}_t$. If: (a) the parameters of these conditional and marginal distributions are not subject to any cross-restrictions; and (b) the parameters of interest can be uniquely obtained from the corresponding parameters of the conditional submodel alone (so that all the parameters of the marginal submodel are nuisance parameters), then $\underline{Z}_t$ is weakly exogenous and $\underline{X}_t$ is endogenous. In this case, $\underline{Z}_t$ may be thought of as *a priori* or fixed from outside the conditional model for purposes of inference on a set of parameters of interest with no loss of information. Hence, the stipulation of weak exogeneity on the choice of the parameters is not, by itself, directly testable.

Weak exogeneity permits efficient estimation of the parameters of interest at the estimation stage and justifies ignoring the marginal stochastic structure generating excluded known exogenous variables $\underline{Z}$'s. To validate projections of included unknown endogenous variables $\underline{X}$'s conditional on a unidirectional flow of information

from the future paths of $\underline{Z}$'s simultaneously at the prediction stage, the concept of strong exogeneity is required. This entails the absence of lagged feedback from the $\underline{X}$'s on to the $\underline{Z}$'s, in addition to weak exogeneity. The cut-off of the linear lagged feedbacks is equivalent to the statement that $\underline{Z}_t$ is not Granger-caused by $\underline{X}_t$ in the system, and is directly testable from relevant sample information. Such a condition for the strong exogeneity of $\underline{Z}$'s guarantees that not only current but lagged values of $\underline{X}_t$ do not affect $\underline{Z}_t$.[12]

For statistical testing purposes, it is customary to use the structural VAR of the form (2.2) with a diagonal white-noise covariance matrix $\Sigma_{\epsilon}=\text{diag}(\sigma_1^2,\ldots,\sigma_K^2)$, $\underline{D}=0$, and $\underline{B}_p(L)=\underline{B}_0-\sum_{i=1}^{p}\underline{B}_i L^i$, but $\underline{B}_0\neq I_K$, to represent the relationship between contemporaneous endogenous and exogenous variables. Suppose now that a subset of variables in the model is exogenously determined by a convincing theory or prior information and that the K-dimensional vector $\underline{Y}_t$ is partitioned accordingly into the M-dimensional endogenous and the (K-M)-dimensional exogenous subvectors $\underline{X}_t$ and $\underline{Z}_t$, respectively. Then the model can be partitioned conformably as follows

$$\underline{B}_p(L)\underline{Y}_t = \begin{bmatrix} \underline{B}_{11}(L) & \underline{B}_{12}(L) \\ \underline{B}_{21}(L) & \underline{B}_{22}(L) \end{bmatrix} \begin{bmatrix} \underline{X}_t \\ \underline{Z}_t \end{bmatrix} = \begin{bmatrix} \underline{\epsilon}_{1t} \\ \underline{\epsilon}_{2t} \end{bmatrix}.$$

Written without the lag operator, we have

$$\begin{bmatrix} \underline{B}_{11,0} & \underline{B}_{12,0} \\ \underline{B}_{21,0} & \underline{B}_{22,0} \end{bmatrix} \begin{bmatrix} \underline{X}_t \\ \underline{Z}_t \end{bmatrix} - \begin{bmatrix} \underline{B}_{11,1} & \underline{B}_{12,1} \\ \underline{B}_{21,1} & \underline{B}_{22,1} \end{bmatrix} \begin{bmatrix} \underline{X}_{t-1} \\ \underline{Z}_{t-1} \end{bmatrix} - \ldots$$
$$- \begin{bmatrix} \underline{B}_{11,p} & \underline{B}_{12,p} \\ \underline{B}_{21,p} & \underline{B}_{22,p} \end{bmatrix} \begin{bmatrix} \underline{X}_{t-p} \\ \underline{Z}_{t-p} \end{bmatrix} = \begin{bmatrix} \underline{\epsilon}_{1t} \\ \underline{\epsilon}_{2t} \end{bmatrix},$$

where $\underline{\epsilon}_{1t}$ and $\underline{\epsilon}_{2t}$ are assumed to be white noise innovations.

---

[12] However, it is essential to realize that $\underline{Z}$'s may not be exogenous to $\underline{X}$'s even though the $\underline{X}$'s do not Granger-cause the $\underline{Z}$'s, since Granger noncausality (to be discussed below) does not rule out the possibility of $\underline{X}_t$ having a contemporaneous effect on $\underline{Z}_t$. [See Enders (1995), pp.315-6, and Granger (1969).]

The assumption that the variables in $\underline{Z}_t$ can be treated as valid conditioning, or exogenous, variables implies that

$$\underline{B}_{21,0} \equiv 0; \quad \text{for weak exogeneity,} \tag{2.9}$$

or

$$\underline{B}_{21,i} \equiv 0; \quad i = 0, \ldots, p, \quad \text{for strong exogeneity.} \tag{2.10}$$

For the time being, suppose the condition of strong exogeneity $(\underline{B}_{21}(L)\equiv 0)$ is met. The parsimonious system then becomes

$$\underline{B}_{11}(L)\underline{X}_t + \underline{B}_{12}(L)\underline{Z}_t = \underline{\epsilon}_{1t}$$
$$\underline{B}_{22}(L)\underline{Z}_t = \underline{\epsilon}_{2t}.$$

The block zero restrictions imposed fulfil the condition:

$$(K-M)M(i+1) \geq \frac{1}{2}K(K-1),^{13}$$

where $K>M\geq 1$. Given $i=p(\geq 1)$ for the null hypothesis of strong exogeneity, the minimum number of exogenous variables (M) for the structural K-dimensional VAR(p) process is chosen from:

$$\frac{1}{2}\left(K-\sqrt{\frac{K^2(p-1)+2K}{p+1}}\right) \leq M \leq \frac{1}{2}\left(K+\sqrt{\frac{K^2(p-1)+2K}{p+1}}\right). \tag{2.11}$$

2.2.2b *Granger and Instantaneous Causality*

In order to examine the relationships between the causes (observed inputs to a system) and the effects (observed outputs from a system), an operational concept of causality has been developed in the econometrics literature. This is that of Granger-causation proposed by Granger (1969), which uses only observed variables for statistical testing. The basic idea is that the cause cannot occur after the effect (temporal priority assumption) and that the cause contains special information about the effect (information

[13] In essence, the constraints imposed must be equal or greater than $[K(K-1)+K]-\frac{1}{2}K(K+1) \geq \frac{1}{2}K(K-1)$, i.e. the difference between the number of parameters in the primitive system and the number of parameters recovered from the standard VAR estimation.

uniqueness assumption). Hence, the arrow of time can be used to help provide a convenient asymmetry between cause and effect, whilst the information in the former can be used to help improve the forecasts of the latter. A useful introduction and discussion can be found in Holland (1986). See also Sims (1982) and references therein.

Suppose that the K-dimensional, weakly stationary process $\underline{Y}_t$ is partitioned into M- and (K-M)-dimensional subprocesses $\underline{X}_t$ and $\underline{Z}_t$ as before, and that the corresponding VAR of formula (2.7) is partitioned accordingly, reflecting an interest in possible causal links between $\underline{X}$ and $\underline{Z}$. The operational statistical testing procedure is then conducted as follows: (a) test whether or not unidirectional causality, in its limited sense,[14] exists from one group of variables $\underline{Z}_t$ to the remaining variables $\underline{X}_t$, without feedback, on the basis of forecasting ability; and (b), if it does exist, then test whether such causality, running from $\underline{Z}_t$ to $\underline{X}_t$, is 'big' or 'small' in the light of the logarithmic ratios of conditional variances suggested by Geweke (1982a).

(1) Tests of the existence of Granger-causality

Consider a K-dimensional, conformably partitioned VAR(p) process $\underline{Y}_t=(\underline{X}_t',\underline{Z}_t')'$ with nonsingular innovation covariance matrix Σ_u,

$$\underline{Y}_t = \begin{bmatrix}\underline{X}_t\\ \underline{Z}_t\end{bmatrix} = \begin{bmatrix}\underline{A}_{11,1} & \underline{A}_{12,1}\\ \underline{A}_{21,1} & \underline{A}_{22,1}\end{bmatrix}\begin{bmatrix}\underline{X}_{t-1}\\ \underline{Z}_{t-1}\end{bmatrix} + \ldots + \begin{bmatrix}\underline{A}_{11,p} & \underline{A}_{12,p}\\ \underline{A}_{21,p} & \underline{A}_{22,p}\end{bmatrix}\begin{bmatrix}\underline{X}_{t-p}\\ \underline{Z}_{t-p}\end{bmatrix} + \begin{bmatrix}\underline{u}_{1t}\\ \underline{u}_{2t}\end{bmatrix}. \quad (2.12)$$

It has been shown that, if $\underline{Y}_t$ embraces all relevant information at time t, a necessary and sufficient condition for $\underline{X}_t$ being not Granger-causal for $\underline{Z}_t$ is that $\underline{A}_{21,i}\equiv 0$, i=1,...,p, while a necessary and sufficient condition for there being no instantaneous causality

[14] For the feasibility of testing, Granger-causality test is carried out merely on the realized history of the series, not on any potentially relevant subjective factors.

between $\underline{X}_t$ and $\underline{Z}_t$ is that $E[\underline{u}_{1t}\underline{u}_{2t}']=0$ (orthogonal innovations).[15] As a result, the existence of Granger-causality can be determined by formal testing for the hypothesized linear restrictions on the VAR coefficients with standard techniques.

(2) Measures of the strength of Granger-causality

After identifying the existence of Granger-causality, it is worth trying to determine the strength of Granger-causality in the framework of stationary VAR(p) processes. With this objective in mind, Geweke (1982a) has monotonically transformed the strength of one-way/two-way causality into the degree of the relevant one-way/two-way feedback, and developed a technique for confirming the extent of various kinds of feedback. In so doing, he achieved a measure of linear interdependence that can be described as the sum of the measures of the three forms of linear feedback, or causality.

Suppose that the focus of attention is on the individual components of the K-dimensional, purely nondeterministic, stationary process $\underline{Y}_t=(\underline{X}_t',\underline{Z}_t')'$, which can be reexpressed as $\underline{X}_t=[I_M \quad 0]\underline{Y}_t=\underline{F}\underline{Y}_t$ and $\underline{Z}_t=[0 \quad I_{K-M}]\underline{Y}_t=\underline{P}\underline{Y}_t$, where $\underline{F}$ and $\underline{P}$ are M×K and (K-M)×K matrices, respectively. Then we have $E[\underline{X}_t]=\underline{F}\mu_Y$, $E[\underline{Z}_t]=\underline{P}\mu_Y$, $\Gamma_X(s)=\underline{F}\Gamma_Y\underline{F}'$ and $\Gamma_Z(s)=\underline{P}\Gamma_Y(s)\underline{P}'$. Hence, $\underline{X}_t$ and $\underline{Z}_t$ are both stationary, purely nondeterministic processes. Application of Wold's theorem implies that, under quite general conditions, $\underline{Y}_t$, $\underline{X}_t$ and $\underline{Z}_t$ each possess finite order VAR representations [see Geweke (1982a), p.305, and Lütkepohl (1991a), p.20, for details]. We have

$$\underline{X}_t = \sum_{i=1}^{m} \underline{C}_{1i}\underline{X}_{t-i} + \underline{v}_{1t}, \quad \text{var}(\underline{v}_{1t})=\Sigma_1, \tag{2.13}$$

or

$$\underline{Z}_t = \sum_{i=1}^{n} \underline{E}_{1i}\underline{Z}_{t-i} + \underline{v}_{2t}, \quad \text{var}(\underline{v}_{2t})=\Sigma_2. \tag{2.14}$$

This system can be treated as the linear projections of $\underline{X}_t$ on $\{\underline{X}_{t-s}|s>0\}$, and of $\underline{Z}_t$ on $\{\underline{Z}_{t-s}|s>0\}$, within which the two

[15] A discussion of these conditions can be found in Lütkepohl (1991a), pp.35-41.

subprocesses $\underline{X}_t$ and $\underline{Z}_t$ are each made up of two main orthogonal parts: one is its own past, the other being the error vector $\underline{v}_{1t}$ or $\underline{v}_{2t}$. Although these errors are each serially uncorrelated, they may be correlated with each other contemporaneously and at various leads or lags.

Suppose now that the focus of attention is shifted to the corresponding joint autoregressive representation (2.12), which can be rewritten as

$$\underline{X}_t = \sum_{i=1}^{p} \underline{A}_{11,i}\underline{X}_{t-i} + \sum_{i=1}^{p} \underline{A}_{12,i}\underline{Z}_{t-i} + \underline{u}_{1t}, \qquad (2.15)$$

and

$$\underline{Z}_t = \sum_{i=1}^{p} \underline{A}_{21,i}\underline{X}_{t-i} + \sum_{i=1}^{p} \underline{A}_{22,i}\underline{Z}_{t-i} + \underline{u}_{2t}, \qquad (2.16)$$

where $(\underline{u}_{1t}', \underline{u}_{2t}')'$ is a white noise process with zero-mean vector and variance-covariance matrix

$$\text{var}\begin{bmatrix} \underline{u}_{1t} \\ \underline{u}_{2t} \end{bmatrix} = E\left(\begin{bmatrix} \underline{u}_{1t} \\ \underline{u}_{2t} \end{bmatrix}\begin{bmatrix} \underline{u}_{1t}' & \underline{u}_{2t}' \end{bmatrix}\right) = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} \equiv \Sigma_u.$$

This system can be viewed as the linear projections of $\underline{X}_t$ and $\underline{Z}_t$ on the same information set $\{\underline{X}_{t-s} | s>0\} \cup \{\underline{Z}_{t-s} | s>0\}$, within which the two component subsystems are mutually statistically dependent with the error vectors $\underline{u}_{1t}$ and $\underline{u}_{2t}$, due to the appearance of the lagged terms $\underline{Z}_{t-s}$ in (2.15) and $\underline{X}_{t-s}$ in (2.16), s>0, each serially uncorrelated but correlated contemporaneously with one another.

If the system (2.15)-(2.16) is premultiplied by the matrix

$$\begin{bmatrix} I_M & -\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{12}'\Sigma_{11}^{-1} & I_{K-M} \end{bmatrix}, \qquad (2.17)$$

which diagonalizes the above white-noise variance-covariance matrix into a K×K block-diagonal matrix

$$\begin{bmatrix} \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}' & 0 \\ 0 & \Sigma_{22}-\Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12} \end{bmatrix},$$

then we have another system, isomorphic to (2.15)-(2.16):

$$\underline{X}_t = \sum_{i=1}^{p} \underline{C}_{2i}\underline{X}_{t-i} + \sum_{i=0}^{p} \underline{D}_{2i}\underline{Z}_{t-i} + \underline{w}_{1t}, \qquad (2.18)$$

and

$$\underline{Z}_t = \sum_{i=0}^{p} \underline{E}_{2i}\underline{X}_{t-i} + \sum_{i=1}^{p} \underline{F}_{2i}\underline{Z}_{t-i} + \underline{w}_{2t}. \qquad (2.19)$$

This system can be regarded as the linear projections of $\underline{X}_t$ on $\{\underline{X}_{t-s}|s>0\}\cup\{\underline{Z}_{t-s}|s\geq 0\}$, and of $\underline{Z}_t$ on $\{\underline{X}_{t-s}|s\geq 0\}\cup\{\underline{Z}_{t-s}|s>0\}$, within which the two component subsystems are mutually statistically independent with the error vectors $\underline{w}_{1t}\left(=\underline{u}_{1t}-\Sigma_{12}\Sigma_{22}^{-1}\underline{u}_{2t}\right)$ and $\underline{w}_{2t}\left(=\underline{u}_{2t}-\Sigma_{12}'\Sigma_{11}^{-1}\underline{u}_{1t}\right)$ each serially uncorrelated and also uncorrelated with each other.

Building upon this canonical framework, Geweke defines the linear feedback from $\underline{Z}$ to $\underline{X}$ as $F_{Z\to X}$, the linear feedback from $\underline{X}$ to $\underline{Z}$ as $F_{X\to Z}$, the instantaneous linear feedback between $\underline{X}$ and $\underline{Z}$ as $F_{X\cdot Z}$, and the linear dependence between $\underline{X}$ and $\underline{Z}$ as $F_{X,Z}$. He concentrates on the nature of the residual variances. That is to say, on

$$F_{Z\to X} = \ln\left[|\Sigma_1|/|\Sigma_{11}|\right],$$

$$F_{X\to Z} = \ln\left[|\Sigma_2|/|\Sigma_{22}|\right],$$

$$F_{X\cdot Z} = \ln\left[|\Sigma_{11}||\Sigma_{22}|/|\Sigma_u|\right],$$

and

$$F_{X,Z} = \ln\left[|\Sigma_1||\Sigma_2|/|\Sigma_u|\right].$$

The implicit motivations involved in the various definitions are: (1) The alternative feedback measures are zero only when the relevant causal ordering is absent; (2) they are all scale invariant and thus remain unchanged under non-singular, scale-preserving linear transformations; and (3) an improvement with the measures is that $F_{Z\to X}$ may be considered a monotonic transformation of the strength of causality, rather than the employment of a less informative all-or-nothing testing procedure [see Bladen-Hovell & Zhang (1991), p.10].

The linkage between these measures is given below:

$$F_{X,Z} = F_{Z \to X} + F_{X \to Z} + F_{X \cdot Z}.$$

Thus the measure of a complete two-sided linear dependence (among stationary series) is the sum of the measures of one-sided linear feedback from the first subseries to the second, and vice versa, on top of an additive category of the nonzero measure of instantaneous linear feedback [see Geweke (1982a), pp.306-7].

§2.3 Critique of Exogeneity and Causality

Exogeneity and causality are regularly employed in the construction of empirical macroeconometric models, but they are still beset with various controversies. Some remarks concerning the two notions may be of importance.

Firstly, as regards exogeneity, the basic issue in the debate over identification restrictions in macroeconometric models eventually boils down to the validity of the dichotomy between endogenous and exogenous variables on *a priori* grounds. Whether it is possible to test the exogeneity assertions about any contemporaneous variables is to a great extent bound up with what is actually meant by exogeneity. Ideally, the only exogenous, or policy, variables would be those for which values can be accepted as extraneous to the system or for which values can be controlled with complete certainty. Nevertheless, the concept of statistical exogeneity defined by Engle *et al.* (1983) is not a property of variables *per se*, but rather a property that variables might have for a set of parameters that are of particular interest to an investigator. Indeed, the limitation of the role of theory in providing acceptable restrictions necessary for identification, and the realization of the form of policy through inspecting signals generated from the operation of the real economy will undoubtedly lead to the conclusion that very few variables are truly exogenous without some compromise.

Secondly, as regards causality, the fundamental objection against the concept of Granger-causality concentrates on the credibility of the definition in characterizing the actual ongoings in the system on a forecasting basis [Judge *et al.* (1985), p.667]. Strictly speaking, the Granger-Geweke tests of the causal chains between the variables involved are tests of predictability or informativeness rather than of causality in an acceptable philosophical sense. More recent contributions in the philosophy of science covering both formal and empirical views of causality can be found in Humphreys (ed) (1994).

However, for further interpreting the absence of causal linkages in the sense that variables in one subvector should have no marked quantitative impact on another, Granger-noncausality is probably inadequate. Note that if (2.12) with maintained zero restrictions ($\underline{A}_{21,i} \equiv 0$, i=1,...,p) is premultiplied by the matrix (2.17), then we can derive another quite different physical representation of the same process, in which variations in $\underline{X}_t$ may significantly affect $\underline{Z}_t$ through the term with coefficient $-\Sigma_{12}'\Sigma_{11}^{-1}$ in the second K-M equations. Therefore, the lack of a Granger-causal relationship, in itself, cannot necessarily be represented as lack of a cause-and-effect relationship without certain simplifications.

Moreover, the critical dependence of the stability of a macroeconomic system on the lag between cause and effect means that 'true' instantaneous causation will never occur between some pairs of stocks (measured through systematic sampling at discrete time points) and flows (measured through aggregation over equal time intervals), however short the frequency of observation of the data. Then any observed or apparent instantaneous causality can be explained by either missing common causes, i.e. jointly unobservable causal variables that cause both or all the variables of interest, or highly temporal aggregation, i.e. improper use of aggregated data in some larger time intervals than actual causal lags [Granger (1988), pp.205-8].

Such temporal aggregation may weaken or even disrupt the plausible causal connections. This is due to the loss of information, or the recording of information at a time other than that of occurrence. To be operational and for use in a forecasting context, a less than usual demanding definition of *prima facie* causality has to be implicitly used. The reasons for this are: the reduction of the information set from all relevant information to past and present values of the process under consideration; the replacement of optimal forecasts by optimal linear forecasts; and the usual choice of minimum forecast mean squared error (MSE) as a measure for forecast precision [Judge *et al.* (1985), pp.668-9]. However, it is reasonable to suggest that, although Granger causal testing may be a matter of dispute, it is still valuable in practice. At the very least, causality analyses can be used for promoting understanding of past data and for forecasting as yet unobserved data.

The critique of the validity of often untested exogeneity assumptions from *a priori* theory in current macroeconometric modelling brings us back to the sphere of a fairly loose, unrestricted VAR (UVAR) approach, where all the variables involved are jointly endogenous with merely minimal prior beliefs incorporated.

The UVAR approach was championed by Sims (1980a, 1982) as an alternative to conventional large-scale macroeconometric models, for studying the dynamic interrelationships among important aggregates, and also for tackling the problem of doubtful *a priori* zero restrictions. Sims doubted the wisdom of developing sophisticated structural macroeconometric models preoccupied with simultaneity as well as possibly 'incredible' overidentifying restrictions. His methodology entails nothing more than selecting K (the number of variables in the UVAR) and p (the maximum lag in each equation) without making reference to any distinction of endogeneity and exogeneity. He did so because, in practice, many complicated relationships are simply poorly understood and none of the non-modelled variables are truly exogenous as many applications (unrealistically) assumed, although statements like this are still

hotly debated [see Hafer & Sheehan (1989)]. As argued in Sims (1980a, pp.5-6):

> [Sometimes variables are] *treated as exogenous only because seriously explaining them would require an extensive modelling effort in areas away from the main interests of the model-builders.*

Whilst statistically well-specified UVAR models have acquired a good track record for producing unconditional forecasts [Bladen-Hovell & Zhang (1991), p.4], the main practical disadvantage stems from the common problem of overfitting or overparameterization. The appearance of overparameterization with too many free insignificant coefficients tends to make either large UVAR out-of-sample forecasts based on these point estimates very inaccurate and volatile or small ones overly sensitive to variations in variable choice.

To alleviate overparameterization and similar problems, Litterman (1980, 1986a) and others developed a technique called the restricted or Bayesian VAR (BVAR) method. Their aim was to incorporate prior and sample information through mixed estimation, and to reduce the influence of the data on the coefficients, rather than reduce the number of coefficients directly. The method can also provide useful information on the dynamic properties of the series and improve forecast performance over other types of analyses. This is taken up in what follows.

§2.4 Conclusion

In this chapter, both structural and standard VAR processes and their statistical properties have been reviewed. The concepts of weak and strong exogeneity as well as the testing and measurement of Granger causality have also been introduced and analysed. Throughout the chapter, it is assumed that all the series are weakly stationary, so that standard estimation methods and standard asymptotic theory could be employed. All of these pave the way for the research that follows.

CHAPTER 3

STATIONARY BAYESIAN VECTOR AUTOREGRESSIONS (BVARs) WITH STOCHASTIC PRIOR RESTRICTIONS

§3.1 Introduction

Statistical forecasting procedures make considerable use of prior beliefs in formulating models through either a restricted parameter space or augmented sample information [see, *inter alia*, Fomby *et al.* (1984), Chapter 6]. However, forecasting methods differ with regard to the main sources of priors, how informative priors about the future economy are to be represented, and how much weight is placed on them. The classical approach tends to utilize inflexible or 'hard' shape priors derived mainly from economic theory to increase degrees of freedom by systematically reducing the parameter space, or the number of free parameters, regardless of historical evidence. The Bayesian procedure, in contrast, incorporates flexible priors obtained mainly from statistical regularities rather than economic theory to increase degrees of freedom by effectively supplementing the sample information. See, for details, Doan (1996, Chapter 8) and Shoesmith (1990, p.261). At the same time, the Bayesian procedure permits the data to modify these prior beliefs if the evidence about coefficients is significant.

In this chapter, a systematic procedure for determining and implementing BVAR forecasting models will be expounded. In general, the construction of an appropriate BVAR forecasting model can be a rather formidable practical problem. It is hoped that the iterative specification search process presented here will reduce this complication and make the model more accessible for economic and business forecasting.

§3.2 Specification of a Finite 'Closed' BVAR Model

Throughout the process of macroeconometric modelling, an essential step is the dynamic specification of an adequate model or models to be estimated. See Harvey (1992, p.11) for relevant discussion. In building a Bayesian VAR forecasting model, one must first construct a UVAR model that is then subject to prior restrictions. Such restrictions can be expressed in the form of subjective probabilities about which one of the possible models will forecast best and should be determined at the second stage in a stepwise "VAR-BVAR" specification search procedure. The preliminary set of specification issues associated with the VAR part of the model are conceptually distinct from those associated with its Bayesian part [cf. Spencer (1993), pp.409-15].

To clarify this standard BVAR specification procedure analytically, it may be instructive to begin with the most basic form of a K-dimensional stationary Gaussian $\underline{Y}_t \sim$VAR(p) process given in (2.3);[1] that is,

$$\underline{Y}_t = \underline{C} + \sum_{i=1}^{p} \underline{A}_i \underline{Y}_{t-i} + \underline{u}_t .$$

The available preforecast multiple time series of the $\underline{Y}$ variables are now divided into p given presample $(\underline{Y}_{-p+1}, \ldots, \underline{Y}_0)$ and T fitted sample $(\underline{Y}_1, \ldots, \underline{Y}_T)$ values. Asymptotically, as T goes to infinity, the effect of the initial, or presample, values vanishes.

Then a typical scalar equation of the linear dynamic system (2.3) is

$$\begin{aligned} Y_{jt} &= C_j + (A_{j1,1}Y_{1,t-1} + A_{j1,2}Y_{1,t-2} + \cdots + A_{j1,p}Y_{1,t-p}) + \cdots \\ &\quad + (A_{jK,1}Y_{K,t-1} + A_{jK,2}Y_{K,t-2} + \cdots + A_{jK,p}Y_{K,t-p}) + u_{jt} \\ &= C_j + \sum_{i=1}^{p} A_{j1,i}Y_{1,t-i} + \cdots + \sum_{i=1}^{p} A_{jK,i}Y_{K,t-i} + u_{jt} ; \end{aligned}$$

[1] Once again, we stick to the Gaussian case, which will make the arguments much simpler.

$$= C_j + \sum_{n=1}^{K} \sum_{i=1}^{p} A_{jn,i} Y_{n,t-i} + u_{jt} ;$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \quad j = 1, \ldots, K; \quad t = 1, \ldots, T. \qquad (3.1)$$

In Eq.(3.1), j indexes the equation number, n indexes the variable number, and i indexes the lag length. Within this framework, there are a total of 1+Kp free coefficients in each equation and a total of $K+K^2p$ free coefficients in the whole system, excluding the parameters of the covariance matrix of the disturbance term.

For T observations, Eq.(3.1) can be written in the compact form:

$$\underline{Y}_j = \underline{X}\underline{\beta}_j + \underline{u}_j; \quad j = 1, \ldots, K, \qquad (3.2)$$

where $\underline{Y}_j = (Y_{j1}, \ldots, Y_{jT})'$ is a (T×1) random vector of observations on the given variable specific to the j-th equation;

$\underline{X} = (\underline{X}_0, \ldots, \underline{X}_{T-1})'$ with $\underline{X}_t = \begin{bmatrix} 1 \\ Y_{1,t} \\ \vdots \\ Y_{1,t-p+1} \\ \vdots \\ Y_{K,t} \\ \vdots \\ Y_{K,t-p+1} \end{bmatrix}$, a (T×(1+Kp)) known common matrix for each equation;

$\underline{\beta}_j = (C_j, A_{j1,1}, \ldots, A_{j1,p}, \ldots, A_{jK,1}, \ldots, A_{jK,p})'$ a ((1+Kp)×1) unknown parameter vector to be estimated; and

$\underline{u}_j = (u_{j1}, \ldots, u_{jT})'$ a (T×1) unobservable disturbance vector with the properties $\underline{u}_j \sim NID(0, \sigma_{jj}^2 I_T)$.

Or more compactly across equations j=1,...,K

$$\underline{Y} = (I_K \circledast \underline{X})\underline{\beta} + \underline{U}, \qquad (3.3)$$

where $\underline{Y} = (Y_{1,1}, \ldots, Y_{1,T}, \ldots, Y_{K,1}, \ldots, Y_{K,T})'$;
(KT×1)

---

[2] The ⊛ sign denotes the Kronecker matrix product. For example, a typical block element of $\underline{A} \circledast \underline{B}$ is $A_{ij} \cdot \underline{B}$.

$$\underset{(T\times(1+Kp))}{\underline{X}}=\begin{bmatrix}1 & Y_{1,0} & \cdots & Y_{1,1-p} & \cdots & Y_{K,0} & \cdots & Y_{K,1-p}\\ \vdots & & & & & & & \vdots \\ 1 & Y_{1,T-1} & \cdots & Y_{1,T-p} & \cdots & Y_{K,T-1} & \cdots & Y_{K,T-p}\end{bmatrix};$$

$$\underline{\beta}=(\underline{\beta}_1',\ldots,\underline{\beta}_K')'$$

$$=(C_1,A_{11,1},\ldots,A_{11,p},\ldots,A_{1K,1},\ldots,A_{1K,p},$$
$$\ldots,$$
$$C_K,A_{K1,1},\ldots,A_{K1,p},\ldots,A_{KK,1},\ldots,A_{KK,p})';$$
$$(K(1+Kp)\times 1)$$

$$\underset{(KT\times 1)}{\underline{U}}=(u_{1,1},\ldots,u_{1,T},\ldots,u_{K,1},\ldots,u_{K,T})';\ \text{and}$$

$$\underline{U}\sim NID(0,\underset{(KT\times KT)}{\Sigma_u\otimes I_T}),\ \text{with the white noise covariance matrix}$$

$$\Sigma_u\equiv\begin{bmatrix}\sigma_{11}^2\cdots\sigma_{1K}\\ \vdots\ \ddots\ \vdots\\ \sigma_{K1}\cdots\sigma_{KK}^2\end{bmatrix}\ \text{assumed known for the moment.}$$

The objective is to construct a suitable BVAR forecasting model with stochastic prior restrictions concerning the VAR coefficients. For this, the two distinct VAR and Bayesian parts of the model must be carefully specified. Construction of an unrestricted VAR usually involves a set of specification issues associated with choices of the set of variables in the VAR and the length of the lag in each equation. The standard Bayesian informative priors, however, explicitly throw light on each variable's own recent lags together with only the most recent lags of other variables in the present confines of vector autoregressions. The Bayesian VAR therefore addresses an additional set of specification issues associated with choices of intercept term and prior structure of the system in a Bayesian mean-variance framework. These specification issues are now discussed.

3.2.1 Determination of VAR Variables and Lag Length

In the selection of the number of variables as well as the number of lags used with these variables, a major goal to achieve is

parsimony. Whether a subset of lagged explanatory regressors should be introduced into, or eliminated from, a given model depends in practice largely upon whether or not it has marked effects on the prediction of a variable being forecast. If the subset contributes little to prediction, it may be a good candidate for omission. See, amongst others, Hafer and Sheehan (1989) or Sims (1980a) for an interesting discussion.

3.2.1a *Determination of VAR Variables*

Usually, the variable selection decision is partly determined by which variables the modeller wishes to forecast, and partly guided by some loose notions concerning which currently known aggregates are potentially related to the forecasted variables and might therefore be useful in forecasting them. However, in view of the limitation in the number of observations and a tradeoff between the inclusion of additional predictors and the concomitant imprecision of statistical estimation, it may be necessary to decide which variables to include and which to exclude on grounds of parsimony. This idea does not imply that the economy concerned is a simple process, merely that relatively simple models can provide good approximations to the actual DGP, and that there are groups of highly correlated variables of which these are representatives [Coen *et al.* (1969), p.136].

Generally, BVAR modellers can be thought of as having priors in choosing a group of candidate variables for inclusion in an initial model specification. Data may then be used subsequently to pick the final set of variables, but only from among a list of prime candidates previously picked by the modellers. The practical problems of the variable selection associated with the BVAR model configuration for the major European economies will be further discussed in Chapter 6.

3.2.1b *Selection of the VAR Order*

In a VAR, long lag lengths quickly consume degrees of freedom. If

lag length p is too small the model can be seriously misspecified, whilst if p is too large degrees of freedom are wasted [see Enders (1995), p.313]. In addition to allowing the variables in the system to trace out a dynamic sequence over time, appropriate lag-length selection can help to conserve degrees of freedom by reducing the dimensionality of the parameter space. However, the optimum lag truncation parameter p is never known *a priori* and must be empirically determined at the specification stage.

Hafer and Sheehan (1989) found that there often exists a superiority of shorter-lagged models over longer-lagged alternatives in forecasting accuracy. A suitable sequential testing procedure to determine the smallest possible lag length is the log-likelihood ratio test suggested and used by Sims (1980a).

The LR statistic recommended for testing a truncated VAR(l) against a full VAR(m), l<m, is

$$LR(l|m) = (T-c)\ln\left(|\hat{\Sigma}_u(l)|/|\hat{\Sigma}_u(m)|\right).$$

Here $\hat{\Sigma}_u(\cdot)$ is the ML estimator of the variance/covariance matrix of residuals obtained by LS estimation either from the restricted VAR(l) or from the unrestricted VAR(m). The multiplier correction c is used to improve the small sample properties of the statistic, and is equal to the number of parameters estimated in each of the unrestricted equations. See, for example, Enders (1995), pp.313-5.

The LR statistic is asymptotically distributed as a chi-square (χ^2) with $K^2(m-l)$ degrees of freedom. In the sequence of testing l=m-1,...,1 against the maximum m, the appropriate estimate $\hat{p}(LR)$ of p is chosen to be l when VAR(l-1) is first rejected, i.e. the outcome of LR(l-1|m) is statistically significant, at some predetermined significance level. Any shortening of the lag length will also produce a significant LR statistic.

It is important to keep in mind that the order chosen in this scheme may not be consistent with the specific purpose of the model

constructed. In some cases, the object is not to obtain precise estimates of these orders, but rather to choose a model specification that is expected to generate accurate forecasts. Other selection criteria include: Akaike's (1969, 1971) final prediction error (FPE) criterion; Akaike's (1973, 1974) information criterion (AIC); Schwarz's (1978) Bayesian information criterion (BIC); and Hannan & Quinn's (1979) (HQ) criterion. They are given in Table 3.1.

Table 3.1: A comparison of four chosen criteria for estimating p

General estimation rule	Choose the optimal estimate $\hat{p}$ such that $F(\hat{p}) = \min\{F(l) \mid l = 0,1,\ldots,m\}$	
Criteria	Objective function	Comparison
FPE	$FPE(l)=\left(\frac{T+Kl+1}{T-Kl-1}\right)^K \lvert\hat{\Sigma}_u(l)\rvert$	$\hat{p}(FPE)=\hat{p}(AIC)$ for all T;
AIC	$AIC(l)=\ln\lvert\hat{\Sigma}_u(l)\rvert+\frac{2}{T}K^2l$	$\hat{p}(AIC)\geq\hat{p}(BIC)$ if $T\geq 8$;
BIC	$BIC(l)=\ln\lvert\hat{\Sigma}_u(l)\rvert+\frac{\ln T}{T}K^2l$	$\hat{p}(AIC)\geq\hat{p}(HQ)$ if $T\geq 16$;
HQ	$HQ(l)=\ln\lvert\hat{\Sigma}_u(l)\rvert+\frac{2\ln\ln T}{T}K^2l$	$\hat{p}(HQ)\geq\hat{p}(BIC)$ for all T.

Note: further references on the above criteria can be found in Judge et al. (1985) and Lutkepohl (1991a).

The asymptotically equivalent FPE and AIC criteria may have the advantage that, asymptotically, the chosen model is never too small, but they have to be judged by a trade-off between decreased bias and increased variance in the coefficient estimates and tend to make the resulting estimates inefficient. The consistent BIC and HQ criteria place relatively more weight on efficiency and choose the correct order, if it exists, asymptotically more often than the former two criteria, but they tend to underfit the chosen model in small or moderate samples. In this study, the likelihood ratio (LR) test statistic modified by Sims (1980a) is employed to select lag length. The LR testing procedures can easily be carried out using the RATS software package.

Litterman (1986a) argues that statistical formulae like those above

are inappropriate for tailoring lag lengths in a VAR model. The more important and appropriate question is, perhaps, how to define an ideal weighted lag scheme so as to express more realistically our true state of knowledge and so to ensure the correct parameter space of our model. Litterman asserts (1986a, p.27):

> *What such formulas ignore is that the reason one wants to choose a lag length in the first place is because one has prior information that more recent values of the variable in question have more information than now distant values.*

Relying primarily on this idea, he consequently proposes one typical BVAR model with a suitable downweighting lag prior on the coefficients. The proposition reflects the fact that coefficients on longer lags are more likely to be close to zero, i.e. that proper lag length truncation, with declining weights, is desirable. Although this alternative approach does not alter the lag lengths *per se*, it does influence the weights placed on lags by way of the lag decay imposed by the modelbuilder's prior information.

Once a vector content and its lag length have been specified for a pth-order, K-variable VAR, the next step is to formulate a multivariate stochastic prior structure for the system.

3.2.2 Determination of Constant Term and Prior Structure

The virtue of Bayesian analysis is that uncertainty about the values of unknown parameters in the model can be formalized in terms of prior distributions which can be defined by certain prior parameters. Moreover, uncertainty about the values of prior parameters can be addressed by assigning to these parameters a further layer of prior distributions. These can be expressed via some scale factors called prior hyperparameters.[3]

A practical advantage of an explicitly Bayesian approach is that it

[3] For some work in this area see Goel & DeGroot (1981), Judge *et al.* (1985), Chapter 4, Lindley & Smith (1972), Smith (1980), Trader (1983), and Trivedi (1980)

provides forecasters with a formal framework for reducing the amount of undocumented or unreproducible 'art' that creeps into a model's construction. It enables them to represent realistically their knowledge, and to effectively combine those beliefs with the historical record.

3.2.2a *Treatment of the Constant Terms*

In practice, with one exception, the standard Bayesian statistical procedure of specifying the same informative, but reasonably diffuse, prior for the parameters of interest, contained in a subset of the parameter vector $\underline{\beta}_j$ in (3.2) proceeds in two main stages. The first is to determine the key features of the prior coefficient probabilities via the best guesses and restrictions that group the own-lag and cross-lag prior standard deviations and govern the relative sizes of the standard deviations within each group. The second is to select, for each of the groups, a suitable value of a hyperparameter that converts the relative sizes of the prior standard deviations to their absolute magnitudes. The only exception to this two-stage procedure concerns an improper flat-prior, by convention, on the intercept in each equation, reflecting our lack of information about its prior distribution.

Our lack of knowledge about prior mean and variance of an intercept implies that all possible values of the constant term have to be treated as (almost) equally likely and to be determined by the data alone. In this sense, a nonzero constant may be either included unrestrictedly in all of the equations, as in (3.1), or use made of a mean-corrected model. Allowing for the removal of the constant term indicates that no prior information is available for the sample mean and that a Bayesian prior can be placed symmetrically on the mean-adjusted process.

3.2.2b *Refinement of the Joint Priors*

Intuitively, prior information is some stock of knowledge, suggested by past experience, which is relevant to some objective of the

modeller. Since the forecasting equations constructed here are dynamic reduced forms, not structural relations, the prior information needed for the VAR coefficients is not derived from theoretical arguments, but from statistical considerations. Additionally, since we do have to assume a specific form for the joint probability distribution, a multivariate normal distribution is almost invariably the form that is chosen for the parameters.

(1) The general multivariate normal priors

In general, a vector of the unknown VAR coefficients $\underline{\beta}$ is jointly normally distributed with prior mean vector $\underline{\beta}^*$ and covariance matrix $\underline{V}_\beta$ if

$$p(\underline{\beta}) = \left(\frac{1}{2\pi}\right)^{K(1+Kp)/2} |\underline{V}_\beta|^{-1/2} \exp\left[-\frac{1}{2}(\underline{\beta}-\underline{\beta}^*)' \underline{V}_\beta^{-1}(\underline{\beta}-\underline{\beta}^*)\right]. \quad (3.4)$$

However, when a noninformative prior is used on the constant terms it may well be preferable to rewrite the joint stochastic prior distribution for each equation of the system in the similar linear form:

$$\underline{r}_j = \underline{R}_j\underline{\beta}_j + \underline{v}_j; \quad j = 1, \ldots, K. \quad (3.5)$$

Here, $\underline{r}_j$ and $\underline{R}_j$ are respectively a (1+Kp)×1 vector and a (1+Kp)×(1+Kp) diagonal matrix of known elements, implying a multivariate normal prior distribution on a subvector of $\underline{\beta}_j$ centred around the chosen values. The vector $\underline{v}_j$ is, with respect to $\underline{r}_j$, a (1+Kp)×1 vector of random errors, assumed to be distributed independently of the $\underline{u}_j$-components with $\underline{v}_j \sim NID(0, I_{1+Kp})$.

A practical advantage of this form is that, in terms of (3.5), an approximate multivariate normal prior can be easily imposed on the parameters of interest (with the constant term left unrestricted). In order to make this kind of prior information useful for the parameter estimates $\hat{\underline{\beta}}_j$ of the j-th equation, $\underline{r}_j$ and $\underline{R}_j$ must be further specified within a formal Bayesian mean-variance framework.

(2) The Minnesota or Litterman priors

When working with a large BVAR forecasting model with numerous coefficients, it is impractical to deliberate about the prior distribution for each of the coefficients individually. In this respect, a simple but useful approach is to use the Minnesota or Litterman prior outlined by Todd (1984). This will specify a full set of prior distributions routinely once the modeller has chosen some of their key features. One feature of this prior is that the meaning of, and linkage between, the values of different coefficients are rather vague or diffuse. Hence, the prior distributions for all stochastic coefficients are assumed to be independent normal so that they can be fully described by their first two moments. The prior covariances are always set to be zero. Moreover, the Minnesota prior allows us to indicate the prior means and variances for the parameters, except the constant term, with reference to a set of prior statistical considerations. Scope remains available to adjust settings for the hyperparameters of the priors in light of the evidence.

(a) The prior mean

With the Minnesota prior, it is taken that the intertemporal dependencies of the modelled variables are likely to be weak and that most of the variation in a given variable is accounted for by its own past. Accordingly, by default, the 'guesstimated' means of the prior distributions for all coefficients of a stationary VAR are usually set to be zero with nonzero prior variances. In other words, $\underline{r}_j=0$ and $\underline{R}_j\neq 0$, $j=1,\ldots,K$. If these guesses were right, each normalized variable would behave very much like a white noise process, around a constant term, in the stationary case.

Given the means of zero, all the modeller would need to do for the rest of the prior is to characterise his degree of confidence in terms of the prior standard deviations, with smaller values reflecting greater confidence in the chosen means.

(b) The prior standard deviation

In the Minnesota prior, it is also believed that the prior standard deviation decreases with increasing lag length. The less recent the value of a variable, the less important it is for forecasting and the greater is the modeller's confidence in the prior value of its coefficient. Moreover, regardless of the different scales of the data series, the prior standard deviations on lag coefficients of the own variable are relatively larger than the corresponding ones of other variables in any given equation.

A two-stage procedure will be adopted in specifying prior standard deviations. The first is to define K non-overlapping groups, for own-lag and all other K-1 sets of cross-lag coefficients, in each equation to determine, within each group, how the coefficients' standard deviations are related to each other. The second is to assign a possible setting of a limited number of hyperparameters, one for each group, to the various groups of coefficients to complete the determination of the prior standard deviations and thus of the prior variances.

Relative standard deviations

The initial relative levels of the standard deviations of the coefficients in the j-th equation of the system can be specified according to two factors. The first factor comes from a lag tightness function $g(l)=1/l$,[4] which implies that as lag length increases, the modeller becomes increasingly confident that coefficients on longer lags will be close to zero, while coefficients on shorter lags are more likely to be nonzero. All the relative weighting patterns imposed upon the prior standard

[4]Two types of lag functions suggest themselves: harmonic ($g(l)=l^{-\gamma}$, $\gamma\geq 0$) and geometric ($g(l)=\gamma^{l-1}$, $0\leq\gamma\leq 1$). Litterman (1986a) advises the use of the harmonic (with the decay parameter $\gamma=1.0$) rather than the geometric function, since the latter seems to get too tight too fast. For a more detailed discussion see Spencer (1993), p.413.

deviations of the coefficients of either own or cross lag variables in each of K groups are of the same harmonic decreasing form. The second factor results from an own-versus-cross standard deviation ratio σ_j/σ_k, which recognizes that the variables in the model are likely to be of different magnitudes, and thus should be adjusted by the ratio to make units comparable. The original relative sizes of overall standard deviations are now replaced by the relative sizes of error standard deviations.[5] This is due to the central assumption that the scale of the response of one variable to another is mainly captured by the unexpected movements contained in the error standard deviations [see Lütkepohl (1991a), pp.209-11].

Once this has been done, it simply remains to pick, for each group of coefficients, just one free hyperparameter that converts all the weights attached to coefficients in the group from relative into absolute prior standard deviations.

Absolute standard deviations

The final absolute levels of the prior standard deviations of the coefficients in the j-th equation of the system can also be specified by two factors. One comes from the overall tightness, or weight, parameter ν_j, which determines the fundamental prior standard deviation on the first own lag. The other results from the relative tightness function f(j,k), which controls the weight on all other variables relative to that on the own variable in each equation.

Weight assignment and adjustment in the scalar ν_j as well as the matrix $\underline{F}=\{f(j,k)\}$ is critical in the specification of prior structure, allowing some elasticity, so that the setting of the hyperparameters could be revised up during the estimation stage. The

[5]Here, σ_j and σ_k are the standard deviations of the errors in an unrestricted univariate autoregression on equations j and k, respectively, and the ratios of σ_j over σ_k, $k\neq j$, are each independent of lag length but specific to the given cross variable k.

Minnesota prior, however, takes advantage of this feature and thus simplifies the standard Bayesian approach by specifying a particular range of values for the relatively few free hyperparameters. All values within the range are treated as equally likely but all values outside it highly unlikely. There are two basic types of priors available: those that correspond to symmetric $\underline{F}$ and general $\underline{F}$.

For a small system with, say, five or fewer equations, a tighter symmetric prior is often preferred [see Doan (1996), p.8-19], which gives the same one relative weight ($w \in [0,1]$) applied to all the off-diagonal variables in the system:

$$f(j,\ k) = \begin{cases} 1.0 & \text{if } j=k \\ w & \text{if } j \neq k. \end{cases} \tag{3.6}$$

Hence

$$\nu\underline{F} = \nu \underset{(K\times K)}{\begin{bmatrix} 1.0 & & w \\ & \ddots & \\ w & & 1.0 \end{bmatrix}} = \underset{(K\times K)}{\begin{bmatrix} \nu & & \nu w \\ & \ddots & \\ \nu w & & \nu \end{bmatrix}},$$

where each element in the rows corresponds to the individual hyperparameter for each group in each equation. Choosing w between 0 and 1 means that, irrespective of differences in scales, other variables would have a smaller prior standard deviation than that of the own variable in relative terms, whereas specifying w=0 implies that the vector system degenerates to a set of univariate autoregressions. In this case, there are only (the same) two free hyperparameters in every equation. The RATS program default for this prior is a simple combination of ν=0.15 and w=0.5.

For a large system with six or more equations, a looser general prior is likely to be desired, which puts moderate weight on important variables and low weight on less important ones in the system:

$$f(j,\ k) = \begin{cases} 1.0 & \text{if } j=k \\ w_{jk} & \text{if } j \neq k. \end{cases} \tag{3.7}$$

Hence

$$\underset{(K\times K)}{\underline{\nu F}}=\begin{bmatrix}\nu_1 & & 0\\ & \ddots & \\ & & \nu_K\end{bmatrix}\underset{(K\times K)}{\begin{bmatrix}1.0 & w_{12} & \cdots & w_{1K}\\ w_{j1} & & 1.0 & w_{jK}\\ w_{K1} & \cdots & w_{K,K-1} & 1.0\end{bmatrix}}=\underset{(K\times K)}{\begin{bmatrix}\nu_1 & \nu_1 w_{12} & \cdots & \nu_1 w_{1K}\\ \nu_j w_{j1} & & \nu_j & \nu_j w_{jK}\\ \nu_K w_{K1} & \cdots & \nu_K w_{K,K-1} & \nu_K\end{bmatrix}},$$

where each element in the rows refers to the particular hyperparameter for each group in each equation. Setting ν_j and w_{jk} too large, such as $\nu_j=3$ and $w_{jk}=1$, will effectively eliminate the influence of the prior, while setting ν_j or w_{jK} too small will either force the own lags too close to the prior mean or eliminate the vector part by cutting out too much interaction.[6] In this case, there is a total of K free hyperparameters in every equation and a total of K×K free hyperparameters in the entire system. A UVAR(K,p) model with K(Kp+1) free parameters is therefore reduced to a more informative BVAR(ν,w) model with at most K(K+1) free parameters through the imposition of prior restrictions on the form of the lag coefficients.

In searching for the most appropriate hyperparameter values within a given range, $0<\nu_j\leq 3$ and $0\leq w_{jK}\leq 1$, say, it may be preferable first to choose a few values for these standard prior hyperparameters and then compare the accuracy of the simulated 'out-of-sample' forecasts from the corresponding BVAR models. This can be done either formally by examining a single value formed by a weighted average of Theil's U values, or informally by studying changing patterns in the values thereof. In practice, the method used is informal. Various possible weights for those variables which are believed to be important in predicting the desired variable will be tried, whilst others believed to be less important have weights that are kept small and constant. We select as the best setting of the hyperparameters, the one that leads to the lowest values of the Theil U statistics.

[6] Here, the value of 0 is a logical lower bound for both ν_j and w_{jk}; the value of 3 and 1 are a generous upper bound for ν_j and w_{jk}, respectively [see Spencer (1993), pp.414-8].

(c) The combined effect

The general standard deviation function, s(j,k,l), for the coefficient $A_{jk,l}$, the jk-th element of $\underline{A}_l$ in the system, is, therefore, specified as

$$s(j,k,l) = \nu f(j,k)g(l)\sigma_j/\sigma_k = \begin{cases} \nu_j/l & \text{if } j=k \\ \nu_j w_{jk}\sigma_j/l\sigma_k & \text{if } j\neq k. \end{cases} \qquad (3.8)$$

In matrix form, we have

$$\underline{S}=\{s(j,k,l)\}=\begin{bmatrix} \nu_1/l & \nu_1 w_{12}\sigma_1/l\sigma_2 & \cdots & \nu_1 w_{1K}\sigma_1/l\sigma_K \\ \vdots & \ddots & & \vdots \\ \nu_j w_{j1}\sigma_j/l\sigma_1 & \nu_j/l & \ddots & \nu_j w_{jK}\sigma_j/l\sigma_K \\ \vdots & & \ddots & \vdots \\ \nu_K w_{K1}\sigma_K/l\sigma_1 & \cdots & \nu_K w_{K,K-1}\sigma_K/l\sigma_{K-1} & \nu_K/l \end{bmatrix}_{(K\times K)}.$$

We now reintroduce the constant terms as the symmetric prior for the other coefficients have been discussed. The typical j-th equation (3.1) with the final prior standard deviations imposed upon its K-group coefficients can be written as

$$\begin{aligned} Y_{jt} = {} & \underset{(\infty)}{C_j} + \underset{(\nu_j w_{j1}\sigma_j/\sigma_1)}{A_{j1,1}Y_{1,t-1}} + \cdots + \underset{(\nu_j w_{j1}\sigma_j/p\sigma_1)}{A_{j1,p}Y_{1,t-p}} + \cdots \\ & + \underset{(\nu_j)}{A_{jj,1}Y_{j,t-1}} + \cdots + \underset{(\nu_j/p)}{A_{jj,p}Y_{j,t-p}} + \cdots \\ & + \underset{(\nu_j w_{jK}\sigma_j/\sigma_K)}{A_{jK,1}Y_{K,t-1}} + \cdots + \underset{(\nu_j w_{jK}\sigma_j/p\sigma_K)}{A_{jK,p}Y_{K,t-p}} + u_{jt}; \end{aligned}$$

$$0<\nu_j\leq 3; \quad 0\leq w_{jk}\leq 1; \quad j,k = 1, \ldots, K; \quad t = 1, \ldots, T, \qquad (3.9)$$

where the prior standard deviations are given in parentheses and an infinite standard deviation is used for the constant term, reflecting the lack of prior information on that coefficient. The corresponding prior covariance matrix $\underline{V}_j$ for the parameter vector of

interest $\underline{\beta}_j$ can be represented as a (1+Kp)×(1+Kp) diagonal matrix

$$\underline{V}_j = \text{diag}\Big[\ \infty,\ (\nu_j w_{j1}\sigma_j/\sigma_1)^2,\ \ldots,\ (\nu_j w_{j1}\sigma_j/p\sigma_1)^2,\ \ldots,$$
$$(\nu_j)^2,\ \ldots,(\nu_j/p)^2,\ \ldots,$$
$$(\nu_j w_{jK}\sigma_j/\sigma_K)^2,\ \ldots,\ (\nu_j w_{jK}\sigma_j/p\sigma_K)^2\Big].$$

Also, in terms of (3.5), an approximate multivariate normal prior distribution for the equation can be readily expressed as a set of stochastic linear combinations of the parameters in $\underline{\beta}_j$

$$\underline{r}_j = \underline{R}_j\underline{\beta}_j + \underline{v}_j; \quad \underline{v}_j \sim \text{NID}\big(0,\ I_{1+Kp}\big);^{7} \quad j = 1,\ \ldots,\ K,$$

where

$$\underline{r}_j = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \underline{R}_j = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \cdot & \sigma_1/\nu_j w_{j1}\sigma_j & & & \cdot \\ \cdot & & \ddots & & \cdot \\ \cdot & & & 1/\nu_j & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & p/\nu_j & \cdot \\ \cdot & & & & \cdot \\ 0 & \cdots & \cdots & p\sigma_K/\nu_j w_{jK}\sigma_j & \end{bmatrix}, \quad \underline{\beta}_j = \begin{bmatrix} C_j \\ A_{j1,1} \\ \vdots \\ A_{jj,1} \\ \vdots \\ A_{jj,p} \\ \vdots \\ A_{jK,p} \end{bmatrix}, \quad \underline{v}_j = \begin{bmatrix} v_{j0} \\ v_{j1,1} \\ \vdots \\ v_{jj,1} \\ \vdots \\ v_{jj,p} \\ \vdots \\ v_{jK,p} \end{bmatrix}.$$

((1+Kp)×1) ((1+Kp)×(1+Kp)) ((1+Kp)×1) ((1+Kp)×1)

This formulation specifies a group of Kp hypothetical distributions around the chosen mean values of each of the coefficients (excepting the constant term). The data can then be examined and combined with the prior to determine a preferred forecasting model. It is the use of this kind of prior information that is intended to produce a useful improvement over conventional techniques.

Once the equations of a tentative BVAR model have been specified in the above four major aspects, the data can be used to estimate

---

[7] That is, $\underline{v}_j$ is standardized so that the standard deviation of each of its elements is one. The identity matrix I_{1+Kp} comes from the assumption that the true standard deviations of the coefficients of the system are just the prior standard deviations as we have defined them and are also constant through time across coefficients, i.e.

$$\text{var}(v_{j0}) = \text{var}(v_{j1,1}) = \ldots = \text{var}(v_{jK,p}) = \text{var}(A_{jj,1}/\nu_j) \equiv 1.$$

coefficients associated with the chosen hyperparameters prior to subsequent model-based forecasting. In addition, an iterative process between misspecification tests and respecification could be repeated until the model is acceptable.

§3.3 Estimation of a Specified BVAR Model

In this section, the discussion will focus upon some special estimating problems associated with the use of a K-variable BVAR with informative priors, for chosen lag length p. We shall explain how the prior information is incorporated to produce operational results in conjunction with a finite sample of data via Bayes' theorem in general and Theil's mixed estimation technique in particular. In effect, inclusion of appropriate prior information augments the available sample size and should improve the precision of the parameter estimates.

3.3.1 The Derivation of a Posterior Distribution via Bayes' Theorem

When both prior and sample information about a vector of unknown parameters are available, they can be combined by Bayes' theorem to form the posterior distribution.[8] Bayes' theorem formalizes the general result that a posterior probability density function (pdf) for the parameters of interest is proportional to a prior pdf times the likelihood function; i.e.,

$$\textit{Posterior pdf} \propto \textit{Prior pdf} \times \textit{Likelihood Function}.$$

In this context, by making $\underline{\beta}$ a parameter vector and $\underline{Y}$ a vector random variable, the theorem can be expressed as

$$p(\underline{\beta}|\underline{Y}) \propto p(\underline{\beta})\, l(\underline{\beta}|\underline{Y}).$$

[8] For this part see, in particular, Griffiths *et al.* (1993), Judge *et al.* (1985), Chapter 4, Lütkepohl (1991a), pp.206-12 and pp.372-5, and Zellner (1985, 1987).

Here $p(\underline{\beta}|\underline{Y})$ is the posterior probability density for $\underline{\beta}$, $p(\underline{\beta})$ is the prior density, summarizing the additional nonsample information about $\underline{\beta}$, and $l(\underline{\beta}|\underline{Y})$ is the likelihood function, summarizing the sample evidence.[9]

A posterior pdf for the VAR coefficients $\underline{\beta}$ can be derived in the manner of Lütkepohl (1991a, pp.206-10). That is, on combining the sample information in the Gaussian Likelihood function

$$l(\underline{\beta}|\underline{Y}) = \left(\frac{1}{2\pi}\right)^{KT/2}|\Sigma_u \otimes I_T|^{-1/2} \times \exp\left[-\frac{1}{2}(\underline{Y}-(I_K\otimes\underline{X})\underline{\beta})'(\Sigma_u^{-1}\otimes I_T)(\underline{Y}-(I_K\otimes\underline{X})\underline{\beta})\right],$$

with the prior information summarized in (3.4), and dropping irrelevant constants from consideration, we obtain the joint posterior density

$$p(\underline{\beta}|\underline{Y}) \propto p(\underline{\beta})l(\underline{\beta}|\underline{Y}) \propto \exp\left\{-\frac{1}{2}\left[(\underline{V}_\beta^{-1/2}(\underline{\beta}-\underline{\beta}^*))'(\underline{V}_\beta^{-1/2}(\underline{\beta}-\underline{\beta}^*)) + ((\Sigma_u^{-1/2}\otimes I_T)\underline{Y}-(\Sigma_u^{-1/2}\otimes\underline{X})\underline{\beta})'((\Sigma_u^{-1/2}\otimes I_T)\underline{Y}-(\Sigma_u^{-1/2}\otimes\underline{X})\underline{\beta})\right]\right\}. \quad (3.10)$$

---

[9] Before proceeding further, a comment concerning the appropriateness of the Bayesian approach is in order. When $\underline{\beta}$ is employed as shorthand for the unknown population parameters, it is often claimed that Bayesians view parameter estimates as random variables, but the view of Rothenberg (1973, p.138) is important here. He states:

> *The Bayesian analysis does not require $\underline{\beta}$ to represent a random outcome of some actual experiment. ... Bayesian decision theory merely argues that people who wish to decide consistently in uncertain situations would act as though $\underline{\beta}$ were a random variable with a certain distribution function.*

Hence, we need to distinguish between the density function for the sampling process where $\underline{\beta}$ is treated as fixed, and the density functions that express our uncertainty about $\underline{\beta}$ and, hence, treat this vector of parameters as random.

Defining

$$\underline{m} = \begin{bmatrix} \underline{V}_\beta^{-1/2}\underline{\beta}^* \\ (\Sigma_u^{-1/2}\otimes I_T)\underline{Y} \end{bmatrix} \text{ and } \underline{M} = \begin{bmatrix} \underline{V}_\beta^{-1/2} \\ \Sigma_u^{-1/2}\otimes\underline{X} \end{bmatrix},$$

the exponent in (3.10) can be rewritten by completing the square on $\underline{\beta}$ as

$$\begin{aligned} -\frac{1}{2}(\underline{m}-\underline{M}\underline{\beta})'(\underline{m}-\underline{M}\underline{\beta}) &= -\frac{1}{2}((\underline{m}-\underline{M}\bar{\underline{\beta}})-\underline{M}(\underline{\beta}-\bar{\underline{\beta}}))'((\underline{m}-\underline{M}\bar{\underline{\beta}})-\underline{M}(\underline{\beta}-\bar{\underline{\beta}})) \\ &= -\frac{1}{2}[(\underline{m}-\underline{M}\bar{\underline{\beta}})'(\underline{m}-\underline{M}\bar{\underline{\beta}})+(\underline{\beta}-\bar{\underline{\beta}})'\underline{M}'\underline{M}(\underline{\beta}-\bar{\underline{\beta}})], \end{aligned} \tag{3.11}$$

where choosing

$$\bar{\underline{\beta}} = (\underline{M}'\underline{M})^{-1}\underline{M}'\underline{m} = [\underline{V}_\beta^{-1}+(\Sigma_u^{-1}\otimes\underline{X}'\underline{X})]^{-1}[\underline{V}_\beta^{-1}\underline{\beta}^*+(\Sigma_u^{-1}\otimes\underline{X}')\underline{Y}], \tag{3.12}$$

the cross-product terms satisfy the following algebraic identity

$$(\underline{\beta}-\bar{\underline{\beta}})'\underline{M}'(\underline{m}-\underline{M}\bar{\underline{\beta}}) = (\underline{\beta}-\bar{\underline{\beta}})'(\underline{M}'\underline{m}-\underline{M}'\underline{M}(\underline{M}'\underline{M})^{-1}\underline{M}'\underline{m}) \equiv 0.$$

Since the first term on the right hand side of (3.11) is independent of $\underline{\beta}$ and may thus be absorbed into the constant of proportionality, we have

$$p(\underline{\beta}|\underline{Y}) \propto \exp[-\frac{1}{2}(\underline{\beta}-\bar{\underline{\beta}})'\bar{\underline{V}}_\beta^{-1}(\underline{\beta}-\bar{\underline{\beta}})], \tag{3.13}$$

where

$$\bar{\underline{V}}_\beta = (\underline{M}'\underline{M})^{-1} = [\underline{V}_\beta^{-1}+(\Sigma_u^{-1}\otimes\underline{X}'\underline{X})]^{-1}. \tag{3.14}$$

With a normal prior and a likelihood function based on Gaussianity, the posterior pdf is a K(1+Kp)-dimensional multivariate normal pdf with posterior mean $\bar{\underline{\beta}}$ and covariance matrix $\bar{\underline{V}}_\beta$; that is, $\underline{\beta}\sim N(\bar{\underline{\beta}},\bar{\underline{V}}_\beta)$.

Of these parameters, $\bar{\underline{\beta}}$ is usually the principal object of inference. According to Eq.(3.12), if $\underline{r}=(\underline{r}_1',\ldots,\underline{r}_K')'=0$ and $\underline{R}=\text{diag}(\underline{R}_1,\ldots,\underline{R}_K)$, substitution of $\underline{\beta}^*$ and $\underline{V}_\beta$ by $\underline{R}^{-1}\underline{r}$ and $(\underline{R}'\underline{R})^{-1}$, respectively, yields

$$\bar{\underline{\beta}} = [\underline{R}'\underline{R} + (\Sigma_u^{-1} \otimes \underline{X}'\underline{X})]^{-1}[\underline{R}'\underline{r} + (\Sigma_u^{-1} \otimes \underline{X}')\underline{Y}]$$
$$= [\underline{R}'\underline{R} + (\Sigma_u^{-1} \otimes \underline{X}'\underline{X})]^{-1}(\Sigma_u^{-1} \otimes \underline{X}')\underline{Y}. \qquad (3.15)$$

In the classical context, a point parameter estimator such as this (i.e. the posterior mean) can be interpreted as a shrinkage estimator and viewed as an optimal point estimator for the parameter vector $\underline{\beta}$, if such exists. The traditional trade-off between decreased bias and increased variance in a Bayesian specification framework disappears in that a convex mean-squared-error loss function is minimized at the posterior mean by including all relevant variables along with prior information that accurately reflects the most likely values of their coefficients [see Litterman (1986a), p.27]. Also, an advantage of this approach is that, by careful choice of a prior distribution, we allow variables, and lags thereof, to enter an equation at the margin, rather than being forced to exercise the extreme choice of inclusion or exclusion.

The same point estimator (the posterior means) for $\underline{\beta}$ can also be attained numerically by using Theil's (1971, pp.347-52) mixed estimation technique. Such a technique involves supplementing the observed data with stochastic prior information concerning the subjective distributions of the coefficients $\underline{\beta}$ and is believed to be capable of greatly increasing the precision of point estimates. The following discusses the technique that was employed here and the problems encountered in parameter estimation.

3.3.2 An Optimal Point Estimator Using Theil's Mixed Estimation Method

We have seen that when the prior information cannot be represented in the form of a proper multivariate normal distribution, it can be preferable to write the information in the form of linear stochastic restrictions and combine it with the sample information through mixed estimation. Given an appropriate interpretation, mixed estimation methods can be regarded as Bayesian estimators. See, for

more details, Fomby *et al.* (1984), Chapter 6, Judge *et al.* (1973), Nagar and Kakwani (1964), Paulus (1975), Swamy and Mehta (1969), Theil (1963, 1971, 1974a, 1974b), and Theil and Goldberger (1961).

The data is assumed to be generated by the model

$$\underset{(T\times 1)}{\underline{Y}_j} = \underset{(T\times(1+Kp))}{\underline{X}} \underset{((1+Kp)\times 1)}{\underline{\beta}_j} + \underset{(T\times 1)}{\underline{u}_j};$$

$$\underline{u}_j \sim NID\left(0, \sigma_{jj}^2 I_T\right); \quad j = 1, \ldots, K,$$

and the prior to follow

$$\underset{((1+Kp)\times 1)}{\underline{r}_j} = \underset{((1+Kp)\times(1+Kp))}{\underline{R}_j} \underset{((1+Kp)\times 1)}{\underline{\beta}_j} + \underset{((1+Kp)\times 1)}{\underline{v}_j};$$

$$\underline{v}_j \sim NID\left(0, I_{1+Kp}\right); \quad j = 1, \ldots, K.$$

The mixed estimator can be obtained by combining these two sources of information in the following partitioned form

$$\underset{\substack{((1+T+Kp)\\ \times 1)}}{\begin{bmatrix} \underline{Y}_j \\ \underline{r}_j \end{bmatrix}} = \underset{\substack{((1+T+Kp)\\ \times(1+Kp))}}{\begin{bmatrix} \underline{X} \\ \underline{R}_j \end{bmatrix}} \underset{\substack{((1+Kp)\\ \times 1)}}{\underline{\beta}_j} + \underset{\substack{((1+T+Kp)\\ \times 1)}}{\begin{bmatrix} \underline{u}_j \\ \underline{v}_j \end{bmatrix}}; \quad j = 1, \ldots, K, \tag{3.16}$$

where the zero mean white-noise processes $\underline{u}_j$ and $\underline{v}_j$ are assumed each to be independent and to be independent of one another. Thus the disturbances $(\underline{u}_j', \underline{v}_j')'$ have a variance-covariance matrix

$$E\left(\begin{bmatrix} \underline{u}_j \\ \underline{v}_j \end{bmatrix}\begin{bmatrix} \underline{u}_j' & \underline{v}_j' \end{bmatrix}\right) = \begin{bmatrix} \sigma_{jj}^2 I_T & 0 \\ 0 & I_{1+Kp} \end{bmatrix}.$$

An application of Aitken's generalized least-squares (GLS) procedure to (3.16) results in the mixed estimator of $\underline{\beta}_j$:

$$\hat{\underline{\beta}}_j = \left(\begin{bmatrix} \underline{X}' & \underline{R}_j' \end{bmatrix}\begin{bmatrix} \sigma_{jj}^2 I_T & 0 \\ 0 & I_{1+Kp} \end{bmatrix}^{-1}\begin{bmatrix} \underline{X} \\ \underline{R}_j \end{bmatrix}\right)^{-1}$$

$$\left(\begin{bmatrix} \underline{X}' & \underline{R}_j' \end{bmatrix}\begin{bmatrix} \sigma_{jj}^2 I_T & 0 \\ 0 & I_{1+Kp} \end{bmatrix}^{-1}\begin{bmatrix} \underline{Y}_j \\ \underline{r}_j \end{bmatrix}\right),$$

or

$$\hat{\underline{\beta}}_j = \left[\underline{R}_j'\underline{R}_j + \sigma_{jj}^{-2}\underline{X}'\underline{X}\right]^{-1}\left[\underline{R}_j'\underline{r}_j + \sigma_{jj}^{-2}\underline{X}'\underline{Y}_j\right]$$

$$= \left[\sigma_{jj}^{2}\underline{R}_j'\underline{R}_j + \underline{X}'\underline{X}\right]^{-1}\underline{X}'\underline{Y}_j; \quad \text{when } \underline{r}_j = 0. \tag{3.17}$$

Moreover, (3.16) can be rewritten in stacked form as

$$\underset{(K(1+T+Kp)\times 1)}{\begin{bmatrix} \underline{Y}_1 \\ \vdots \\ \underline{Y}_K \\ \underline{r}_1 \\ \vdots \\ \underline{r}_K \end{bmatrix}} = \underset{(K(1+T+Kp)\times K(1+Kp))}{\begin{bmatrix} \underline{X} & & 0 \\ & \ddots & \\ 0 & & \underline{X} \\ \underline{R}_1 & & 0 \\ & \ddots & \\ 0 & & \underline{R}_K \end{bmatrix}} \underset{(K(1+Kp)\times 1)}{\begin{bmatrix} \underline{\beta}_1 \\ \vdots \\ \underline{\beta}_K \end{bmatrix}} + \underset{(K(1+T+Kp)\times 1)}{\begin{bmatrix} \underline{u}_1 \\ \vdots \\ \underline{u}_K \\ \underline{v}_1 \\ \vdots \\ \underline{v}_K \end{bmatrix}},$$

or, more compactly, as

$$\begin{bmatrix} \underline{Y} \\ \underline{r} \end{bmatrix} = \begin{bmatrix} I_K \otimes \underline{X} \\ \underline{R} \end{bmatrix}\underline{\beta} + \begin{bmatrix} \underline{U} \\ \underline{V} \end{bmatrix}.$$

Here $\underline{Y}=(\underline{Y}_1',\ldots,\underline{Y}_K')'$, $\underline{r}=(\underline{r}_1',\ldots,\underline{r}_K')'=0$, $\underline{R}=\text{diag}(\underline{R}_1,\ldots,\underline{R}_K)$ (a block diagonal matrix), $\underline{\beta}=(\underline{\beta}_1',\ldots,\underline{\beta}_K')'$, $\underline{U}=(\underline{u}_1',\ldots,\underline{u}_K')'$ and $\underline{V}=(\underline{v}_1',\ldots,\underline{v}_K')'$. The variance-covariance matrix of the disturbances $(\underline{U}',\underline{V}')'$ is

$$\left[\begin{array}{ccc|ccc} \sigma_{11}^{2}I_T & \cdots & \sigma_{1K}I_T & & & \\ & \ddots & & & 0 & \\ \sigma_{K1}I_T & \cdots & \sigma_{KK}^{2}I_T & & & \\ \hline & & & I_{1+Kp} & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & I_{1+Kp} \end{array}\right] = \begin{bmatrix} \Sigma_u \otimes I_T & 0 \\ 0 & I_{K(1+Kp)} \end{bmatrix}.$$

The augmented GLS parameter estimates are

$$\begin{aligned} \hat{\underline{\beta}} &= \left\{\begin{bmatrix} I_K \otimes \underline{X} \\ \underline{R} \end{bmatrix}'\begin{bmatrix} \Sigma_u \otimes I_T & 0 \\ 0 & I_{K(1+Kp)} \end{bmatrix}^{-1}\begin{bmatrix} I_K \otimes \underline{X} \\ \underline{R} \end{bmatrix}\right\}^{-1} \\ &\qquad \left\{\begin{bmatrix} I_K \otimes \underline{X} \\ \underline{R} \end{bmatrix}'\begin{bmatrix} \Sigma_u \otimes I_T & 0 \\ 0 & I_{K(1+Kp)} \end{bmatrix}^{-1}\begin{bmatrix} \underline{Y} \\ \underline{r} \end{bmatrix}\right\} \\ &= \left[\underline{R}'\underline{R} + (\Sigma_u^{-1}\otimes\underline{X}'\underline{X})\right]^{-1}\left[\underline{R}'\underline{r} + (\Sigma_u^{-1}\otimes\underline{X}')\underline{Y}\right] \\ &= \left[\underline{R}'\underline{R} + (\Sigma_u^{-1}\otimes\underline{X}'\underline{X})\right]^{-1}(\Sigma_u^{-1}\otimes\underline{X}')\underline{Y}. \end{aligned} \tag{3.18}$$

By comparing (3.15) and (3.18), it is seen that the results obtained above are exactly in the same form as the posterior means defined previously. Hence, in this simple mixed estimation procedure, the resultant estimated values of $\underline{\beta}$ will just be the posterior means (i.e. $\hat{\underline{\beta}}\equiv\bar{\underline{\beta}}$) and serve as optimal point estimator under the convenient simplifying assumption that priors and disturbances are all normally distributed.

In practice, Σ_u is rarely known and the inversion of the matrix $\underline{R}'\underline{R}+\left(\Sigma_u^{-1}\otimes\underline{X}'\underline{X}\right)$ of dimension $(K(1+Kp)\times K(1+Kp))$ in the above estimator for $\underline{\beta}$ can be intractable. Since BVAR priors enable one to choose a relatively large p, the dimension of the matrix to be inverted in computing $\hat{\underline{\beta}}$ may be quite substantial. To get rid of these problems, (3.17) may be used for each equation of the system individually, with Σ_u replaced by the ML estimator $\tilde{\Sigma}_u$.[10] It should be stressed, however, that point estimators of $\underline{\beta}$, say the posterior mean $(\bar{\underline{\beta}})$, important though they are, are only special applications of the posterior pdf and are generally an inadequate means of reporting results. It is, therefore, more useful to report complete posterior distributions for the parameters.

There are three key points that need to be emphasized. Firstly, the whole point of introducing GLS for the BVAR model is to improve the efficiency of estimation (or to reduce the estimator variances). In a standard multivariate vector autoregression, any regressor which appears on the right hand side of each of the K equations would belong in principle to the right hand side of all of them. In this case, if the informative priors are not imposed on the system, the GLS estimator, either for the system or for the equation, will reduce to the ordinary least squares (OLS) estimator for the UVAR

---

[10]The ML estimator of the white noise covariance matrix Σ_u can be written as

$$\tilde{\Sigma}_u = \underset{(K\times T)}{\underline{Y}}\left(I_T-\underset{((1+Kp)\times T)}{\underline{X}}'\left(\underline{X}\underline{X}'\right)^{-1}\underline{X}\right)\underline{Y}'/T,$$

where $\underline{Y}=[\underline{Y}_1,\ldots,\underline{Y}_T]$, and $\underline{X}=\underline{X}'$.

model. In contrast, if the prior distributions are constrained to specific values, the GLS estimator will reduce to restricted least squares (RLS) for the structural VAR. Thus OLS and RLS are only special cases of mixed estimation.

Secondly, the seemingly unrelated regression (SUR) estimator is not needed for a system that has exactly the same set of variables on the right hand sides and exhibits solely contemporaneous correlations of the disturbances, as there is no efficiency gain from estimating the system as a whole. It is known that in any multiple equation model, the use of restrictions on parameter values tends to reduce estimator variances and the estimator of a complete system would, in general, be no less efficient than the estimator of any other cases. If there were different degrees of restrictions on the regressors in different equations, the existence of some connection between equations would enable a restriction on one equation to improve the efficiency of the estimates of other equations in the system. Only if all equations have exactly the same set of regressors and small residual correlations will no one equation be more restricted than any other, compared to a global list of all regressors in the model. Hence, GLS estimation applied equation by equation would be as efficient as the whole system, and would appear to be much cheaper than the full information method of estimation.

Finally, the difference between GLS estimates (for the system with a prior) and OLS estimates (for the system without a prior) is that, with proper priors, the degrees of freedom reported are no longer T-1-Kp for each equation but T-1 instead. T-1 has no relation with the lag length p, since BVAR priors are typically defined for each right-hand-side variable except for the intercept, and may be much larger than T-1-Kp. This represents a somewhat artificial way of getting around the problem that, with a prior, Kp can exceed T [see Doan (1996), p.8-21]. The degrees of freedom are increased by stochastic restrictions, while the variability of the estimates is reduced and the reduction in overparameterization should yield improved out-of-sample forecasts [see Shoesmith (1990), p.261].

After the coefficients of a BVAR model have been estimated, the adequacy of the fitted model should be assessed by the use of various evaluation and model selection criteria. If the ultimate goal is prediction, then it seems logical to judge the model, and thus the corresponding modelling procedure, by determining whether the quality of within-sample forecasts is better when some kind of prior information is incorporated, as was done in Litterman (1986a, 1986c). Indeed, the validation of a model simply means whether the selected model, with all its inherent imperfections, does an adequate job in postsample prediction. Although validation of a model from historical data can never, in itself, ensure the quality of forecasts into the future, it is reasonable to believe that a model which has shown itself to be valid on the basis of the past and, hence, has presumably captured some constant features of the underlying structure, will yield more reliable forecasts in the short run than one which has not been so validated. Once a satisfactory BVAR forecasting model is found, optimal linear forecasting is a routine affair [Judge *et al.* (1985), p.660].

§3.4 Prediction Using the Estimated BVAR Model

The objective here is to derive a Bayesian predictive distribution, which contains all of the uncertainty about future events and is a function of known quantities only. Suppose that we wish to derive a complete predictive density for future values on the vector $\underline{Y}_t=(Y_{1t},\ldots,Y_{Kt})'$ over h periods, say $\tilde{\underline{Y}}$, together with a set of Bayesian point predictions of these values, which are assumed to link with a known h×(1+Kp) matrix $\tilde{\underline{X}}$ through the same multivariate normal regression model generating the past sample observations $\underline{Y}$. That is,

$$\tilde{\underline{Y}} = (I_K \otimes \tilde{\underline{X}})\underline{\beta} + \tilde{\underline{U}}. \qquad (3.19)$$

Here $\tilde{\underline{Y}}=(Y_{1,T+1},\ldots,Y_{1,T+h},\ldots,Y_{K,T+1},\ldots,Y_{K,T+h})'$ is a Kh×1 vector

of future values to be forecast over the forecast horizon $t=T+1,\ldots,T+h$;

$$\tilde{\underline{X}}=\begin{bmatrix} 1 Y_{1,T} & \cdots & Y_{1,T+1-p} & \cdots & Y_{K,T} & \cdots & Y_{K,T+1-p} \\ & & & \vdots & & & \\ 1 Y_{1,T+h-1} & \cdots & Y_{1,T+h-p} & \cdots & Y_{K,T+h-1} & \cdots & Y_{K,T+h-p} \end{bmatrix}$$ is a $h\times(1+Kp)$ matrix of given or hypothesised values for the predictor variables, assumed known prior to making a forecast on $\underline{Y}_t$;

$\underline{\beta}$ is a $K(1+Kp)\times 1$ vector of regression parameters for $\tilde{\underline{Y}}$, assumed the same as that for the previous observations $\underline{Y}$; and

$\tilde{\underline{U}}=(u_{1,T+1},\ldots,u_{1,T+h},\ldots,u_{K,T+1},\ldots,u_{K,T+h})'$ is a $Kh\times 1$ vector of future disturbance terms, assumed Gaussian as well with $\tilde{\underline{U}}\sim NID(0,\underset{(Kh\times Kh)}{\Sigma_u\otimes I_h})$ and $E(\underline{U}\tilde{\underline{U}}')=0$.[11]

Since the value of $\underline{\beta}$ is unknown, one way of deriving the Bayesian predictive pdf based on (3.21) is to write down the joint pdf $p(\tilde{\underline{Y}},\underline{\beta}|\underline{Y})$, and then integrate with respect to $\underline{\beta}$ to obtain the marginal, or predictive, pdf $f(\tilde{\underline{Y}}|\underline{Y})$ for $\tilde{\underline{Y}}$. In such a procedure, the form of $p(\tilde{\underline{Y}},\underline{\beta}|\underline{Y})$ can be represented by factorising it into $l(\tilde{\underline{Y}}|\underline{\beta},\underline{Y})\cdot p(\underline{\beta}|\underline{Y})$, with $p(\underline{\beta}|\underline{Y})$ being the posterior pdf for the regression coefficients $\underline{\beta}$, shown in (3.15) and $l(\tilde{\underline{Y}}|\underline{\beta},\underline{Y})$ the likelihood function for the future observations $\tilde{\underline{Y}}$, given by

$$\begin{aligned} l(\tilde{\underline{Y}}|\underline{\beta},\underline{Y}) &= l(\tilde{\underline{Y}}|\underline{\beta}) \\ &= \left(\frac{1}{2\pi}\right)^{Kh/2}|\Sigma_u\otimes I_h|^{-1/2} \\ &\quad\times \exp\left[-\frac{1}{2}(\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\underline{\beta})'(\Sigma_u^{-1}\otimes I_h)(\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\underline{\beta})\right], \\ &\quad \text{noting } E(\underline{U}\tilde{\underline{U}}')=0. \end{aligned}$$

Therefore, the predictive pdf for $\tilde{\underline{Y}}$, $f(\tilde{\underline{Y}}|\underline{Y})$, is readily obtained by the following integration:

$$f(\tilde{\underline{Y}}|\underline{Y}) = \int p(\tilde{\underline{Y}},\underline{\beta}|\underline{Y})\,d\underline{\beta}$$

[11] This arises because the knowledge of the realized values of $\underline{U}$ is of no help in predicting $\tilde{\underline{U}}$. They are mutually orthogonal or temporally independent and hence the conditional expectations of future error terms $\tilde{\underline{U}}$ are zero.

$$= \int l(\tilde{\underline{Y}}|\underline{\beta})\, p(\underline{\beta}|\underline{Y})\, d\underline{\beta}$$

$$\propto \int \exp\Big\{-\frac{1}{2}\big[(\bar{\underline{V}}_{\beta}^{-1/2}(\underline{\beta}-\bar{\underline{\beta}}))'(\bar{\underline{V}}_{\beta}^{-1/2}(\underline{\beta}-\bar{\underline{\beta}}))$$

$$+ ((\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{\beta})'((\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{\beta})\big]\Big\} d\underline{\beta}. \tag{3.20}$$

Performing the integration in (3.20), the predictive pdf can be rewritten as

$$f(\tilde{\underline{Y}}|\underline{Y}) \propto |[\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\bar{\underline{\beta}}]'\underline{W}[\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\bar{\underline{\beta}}]|^{-\gamma/2}, \tag{3.21}$$

where

$$\underline{W} = (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'](\Sigma_u^{-1/2}\otimes I_h);$$

$$\underline{N} = \bar{\underline{V}}_{\beta}^{-1}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}}); \text{ and}$$

$$\gamma = T+h-(1+Kp).$$

From (3.21) it follows that $f(\tilde{\underline{Y}}|\underline{Y})$ is in the form of a multivariate Student t distribution with mean vector

$$E(\tilde{\underline{Y}}) = (I_K\otimes\tilde{\underline{X}})\bar{\underline{\beta}},$$

and covariance matrix

$$E(\tilde{\underline{Y}}-E(\tilde{\underline{Y}}))(\tilde{\underline{Y}}-E(\tilde{\underline{Y}}))'$$

$$= E((I_K\otimes\tilde{\underline{X}})\underline{\beta}+\tilde{\underline{U}}-(I_K\otimes\tilde{\underline{X}})\bar{\underline{\beta}})((I_K\otimes\tilde{\underline{X}})\underline{\beta}+\tilde{\underline{U}}-(I_K\otimes\tilde{\underline{X}})\bar{\underline{\beta}})'$$

$$= E((I_K\otimes\tilde{\underline{X}})(\underline{\beta}-\bar{\underline{\beta}})+\tilde{\underline{U}})((I_K\otimes\tilde{\underline{X}})(\underline{\beta}-\bar{\underline{\beta}})+\tilde{\underline{U}})'$$

$$= (I_K\otimes\tilde{\underline{X}})E[(\underline{\beta}-\bar{\underline{\beta}})(\underline{\beta}-\bar{\underline{\beta}})'](I_K\otimes\tilde{\underline{X}})'+E(\tilde{\underline{U}}\tilde{\underline{U}}')$$

$$= (I_K\otimes\tilde{\underline{X}})\bar{\underline{V}}_{\beta}(I_K\otimes\tilde{\underline{X}})'+\Sigma_u\otimes I_h.$$

Here again, Σ_u is assumed known throughout, and $(\underline{\beta}-\bar{\underline{\beta}})$ and $\tilde{\underline{U}}$ are uncorrelated (as $E(\underline{U}\tilde{\underline{U}}')=0$), enabling us to investigate the structure of forecasting error covariances.

[12] See, for details, Appendix B.

Moreover, since

$$\begin{aligned}\underline{W}^{-1} &= (\Sigma_u^{-1/2}\otimes I_h)^{-1}\left[I_{Kh}+(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}})\underline{\bar{V}}_\beta(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}})'\right]\left[(\Sigma_u^{-1/2}\otimes I_h)'\right]^{-1}\\ &= \left[(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes I_h)\right]^{-1}+(I_K\otimes\underline{\tilde{X}})\underline{\bar{V}}_\beta(I_K\otimes\underline{\tilde{X}})',\\ &= \Sigma_u\otimes I_h+(I_K\otimes\underline{\tilde{X}})\underline{\bar{V}}_\beta(I_K\otimes\underline{\tilde{X}})',\end{aligned}$$

we obtain

$$E\left(\underline{\tilde{Y}}-E(\underline{\tilde{Y}})\right)\left(\underline{\tilde{Y}}-E(\underline{\tilde{Y}})\right)' \equiv \underline{W}^{-1}.^{13}$$

These results enable us not only to make predictions about the elements of $\underline{\tilde{Y}}$ conveniently but to incorporate both data and expert judgement adequately into a final Bayesian predictive distribution.

Given a set of predictions, an optimal point prediction $(\underline{\hat{Y}})$ obtained via the minimization of squared prediction error loss is the mean of the predictive distribution, with a predictive expected loss equal to the variance. Hence

$$\begin{aligned}\underline{\hat{Y}} &= E(\underline{\tilde{Y}}) = (I_K\otimes\underline{\tilde{X}})\underline{\bar{\beta}} \equiv (I_K\otimes\underline{\tilde{X}})\underline{\hat{\beta}} \quad \text{since } \underline{\bar{\beta}} \equiv \underline{\hat{\beta}}\\ &= (I_K\otimes\underline{\tilde{X}})\left[\underline{R}'\underline{R}+(\Sigma_u^{-1}\otimes\underline{X}'\underline{X})\right]^{-1}\left[\underline{R}'\underline{r}+(\Sigma_u^{-1}\otimes\underline{X}')\underline{Y}\right], \qquad (3.22)\end{aligned}$$

or more simply

$$\begin{aligned}\underline{\hat{Y}}_j &= E(\underline{\tilde{Y}}_j) = \underline{\tilde{X}}\underline{\bar{\beta}}_j \equiv \underline{\tilde{X}}\underline{\hat{\beta}}_j\\ &\equiv \underline{\tilde{X}}\left[\underline{R}_j'\underline{R}_j+\sigma_{jj}^{-2}\underline{X}'\underline{X}\right]^{-1}\left[\underline{R}_j'\underline{r}_j+\sigma_{jj}^{-2}\underline{X}'\underline{Y}_j\right]; \quad j = 1, \ldots, K \qquad (3.23)\end{aligned}$$

It can be seen from (3.23) that for the j-th equation of the system, the optimal point predictions based on the predictive density function are just the known fixed matrix $\underline{\tilde{X}}$ of future predictor values times the optimal posterior mean estimator $\underline{\bar{\beta}}_j$ (or the mixed estimator $\underline{\hat{\beta}}_j$). These forecasts constitute an important basis for optimal linear ex ante forecasting, and will be compared with actual values. The accuracy of the BVAR model's forecasts will be further compared with those generated by alternative methods, using Theil's inequality index. It is also seen from (3.23) that the stochastic

[13] Again, see Appendix B.

prior restrictions for the model's coefficients would yield a significant influence on the final point predictions of the variables being considered, especially for small sample sizes.

§3.5 Conclusion

This chapter describes how and why we prefer BVARs to UVARs. In particular, we have described carefully four major aspects involved in the construction of a stationary Gaussian BVAR process: the choice of vector content; the selection of lag length; the treatment of constant terms; and finally, the determination of prior structure for all individual coefficients. We have also considered related issues associated with mixed-estimation and ex ante forecasting.

Following from the work of Sims (1980a, 1982) and others, it has been argued that economic theory is often not very explicit about econometric relationships, and the specification of a multiple time series model usually has to rely, at least to some extent, on sample information. In this regard, UVAR models have been proposed as a class of fairly general models, which do not impose rigid *a priori* restrictions on the data generation process and let data speak for themselves. However, since we often work with limited data, the data cannot speak very loudly in standard UVAR models. UVARs involving a large number of parameters and very few restrictions can often lead to serious problems of overparameterization, with many insignificant coefficients hampering the interpretation of the results.

It is the appearance of the large number of parameters in UVARs that points in the direction of using Bayesian methods. The BVAR model provides us with a formal and flexible way of characterizing stochastic subjective information and combining it with sample data. So far, we have only discussed the formulation of a stationary BVAR model. This is obviously not enough, since many series of practical interest are intrinsically nonstationary. In the following chapters, we will consider generalizations to deal with this.

CHAPTER 4

VECTOR AUTOREGRESSIVE ERROR CORRECTION MODELS (VAR-ECMs) WITHOUT A PRIOR

§4.1 Introduction

This chapter discusses the problems arising from the existence of trends in nonstationary macroeconomic time series. After this, the following chapter incorporates this body of information into a more general BVAR model. Of the BVAR forecasting method noted in Chapter 3, all the modelled variables are assumed to be weakly stationary stochastic processes, at least around some deterministic linear time trends which could be removed. This would imply that there are no trends or shifts in the covariances or seasonal patterns, where standard asymptotic distribution theory can be applied. However, there is ample empirical evidence that many macro series are nonstationary. Most nonstationary series are integrated of order unity (denoted I(1)). This is due to the presence of one unit root [see, e.g. the seminal note of Nelson and Plosser (1982)], which gives rise to a stochastic trend in variance, as opposed to a purely deterministic trend in mean, with shocks or innovations to the economic system being permanent rather than transient.

In traditional time series analyses, a valuable device in handling trends and seasonal patterns is that of differencing. A model can then be fitted to the differenced data. However, a simple VAR or BVAR on differences only would not be able to capture appropriately any long-run information among subsets of nonstationary variables. Current practice in dealing with sets of interacting variables is to attempt to analyse nonstationary series directly and to consider the possibility of dominant 'long-swing' components, such as the trend, being common to more than one series. The concept of cointegration, developed out of the existence of unit roots and a common stochastic trend (CST), was first suggested in Granger (1981). Cointegration,

as well as its intimate connection with associated error correction mechanisms (ECMs), has engaged increasing attention over recent years. In particular, it has provided a joint parametric treatment of the short-run dynamics and the long-run relations of the underlying process within an OLS framework.

In this chapter, the complementary roles of theoretical and empirical analyses in macroeconometric modelling will be reexamined. A number of testing, estimation and inference procedures will be reconsidered in the context of VARs with unit roots. In addition, a linear transformation to I(0) space in terms of differences and cointegrating combinations of observed macro data will be given.

§4.2 A Linear Stochastic Multicointegrated System for Integrated Time Series

Granger and Newbold (1974), before the introduction of the concept of cointegration, investigated by Monte Carlo experiment the likely consequences of specifying a static OLS regression of one random walk on another random walk, independent of the first. They pointed out that the usual significance tests performed on the regression coefficients may be very misleading, resulting in the acceptance of a spurious relationship. Warning signs may be an apparently acceptable R^2 coupled with a low Durbin-Watson (DW) statistic, and high significance of the coefficients (a simple rule of thumb being R^2>DW). Granger and Newbold therefore suggested that, in such circumstances, one should look for a model to apply to the differenced series, rather than to the raw series concerned.

Phillips (1986) provided analytical results explaining the Monte Carlo findings reported by Granger and Newbold. His study demonstrates the following. Firstly, in contrast to the stationary case, the regression coefficients in a model with data generated by integrated processes do not converge in probability to constant coefficient matrices. Secondly, because the regression error is I(1) under the null hypothesis of no relationship with undesirable

infinite variance, the conventional t- and F-statistics, formed as ratios of sample, or joint sample, moments do not have limiting distributions. Upon suitable standardisation, they converge weakly to appropriately defined functionals of vector Brownian Motions or Wiener processes, and actually diverge as the sample size $T\rightarrow\infty$. Hence, there are no asymptotically correct critical values for these test procedures: the larger the sample, the greater will be the rejection rate when the tests are based on a given critical value delivered from conventional asymptotics. Finally, under the same conditions, DW converges in probability to zero, whereas R^2 has a nondegenerate limiting distribution. The spurious regression problem therefore becomes even more pronounced with low values for DW and moderate values for R^2 as $T\rightarrow\infty$.

All these theoretical results are consistent with the empirical findings in the Granger-Newbold (1974) experiments. Indeed, it is the appearance of the integrated error term in a regression between the variables that eventually leads the usual asymptotic distribution theory to break down at the unit circle. Achieving a weakly stationary, or I(0), error with finite variance is then a necessary or minimum condition for a specified model to be satisfactory. However, the theory of cointegration addresses the impact of equilibrium relationships implied by economic theories, within a dynamic adjustment process. It may thus improve long-term forecastability of the system under scrutiny. See, *inter alia*, Davidson and Mackinnon (1993), Chapters 19 and 20, and Hamilton (1989).

4.2.1 Integrated/Cointegrated Processes and Their Properties

In empirical studies, it is frequently of interest to test whether a set of macroeconomic variables move together. There is a tacit assumption that certain economic series cannot be expected to drift too far apart, due to economic forces identified in theory, even though the series themselves may wander extensively over time. The idea of cointegration can be used to test the correctness of such beliefs, and incorporate the resulting information in the specification of dynamic models. First, the definitions and

time-series properties of integrated and cointegrated variables will be clarified.

According to Wold's decomposition theorem, a zero-mean purely nondeterministic stationary stochastic process has an infinite MA representation. This can generally be approximated by a finite ARMA process. However, many economic series need to be differenced in order to make them stationary. A K-dimensional vector $\underline{Y}_t=(Y_{1t},\ldots,Y_{Kt})'$ with no deterministic component is said to be integrated of order d, denoted $\underline{Y}_t \sim I(d)$, if it has a stationary, invertible, multivariate ARMA representation after differencing d times. See Engle and Granger (1987, p.252).

Consequently, a vector integrated of order zero is stationary in levels while a vector integrated of order unity is stationary in first differences. The word 'integrated' is thus a description of the original series relative to the resulting stationary series. There are many substantial differences in appearance between series that are I(0) (with some long-run mean) and I(1) (with some starting value), and the sum of an I(0) and an I(1) is always I(1). See, for a more detailed discussion, Engle and Granger (1987, pp.252-3).

Suppose now that the series of interest in a bivariate vector $\underline{Y}_t$ are all I(d), having no drifts or trends in mean. Then it is generally true that any linear combination of these two series is also I(d). However, in contrast to the spurious regression case, it is possible that there exists some nonzero 2-vector $\underline{\alpha}=(\alpha_1,\alpha_2)'$, such that the combination

$$\underline{\alpha}'\underline{Y}_t = [\,\alpha_1\ \alpha_2\,]\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \sum_{i=1}^{2}\alpha_i Y_{it} = \varepsilon_t \qquad (4.1)$$

is integrated of a lower order, say I(d-b), $d \geq b > 0$. When this occurs, a very special linear constraint operates on the long-run components of the series. Suppose further that d=b=1, so that each variable is I(1), yet ε_t is I(0), i.e. is stationary. Hence the $\underline{Y}$ variables must have long-run components which effectively cancel out to produce ε_t. In such circumstances, the integrated variables under examination

are termed cointegrated. To formalize these ideas, Engle and Granger (1987, p.253) provide the following definition of cointegration.

The components in a K-dimensional vector $\underline{Y}_t$ are said to be cointegrated of order (d,b), denoted $\underline{Y}_t \sim CI(d,b)$, if: (i) all components of $\underline{Y}_t$ are I(d); and (ii) there exists at least one vector of weights $\underline{\alpha}(\neq 0)$ such that $\varepsilon_t = \underline{\alpha}'\underline{Y}_t \sim I(d-b)$, $b>0$. Such a linear combination is a cointegrating combination and the scaling vector $\underline{\alpha}$ which represents it is a cointegrating vector (CV).

Three, among a number of, important points to note about the definition are: (a) cointegration refers to a linear combination of nonstationary variables; (b) all variables must be integrated of the same order; and (c) most of the cointegration literature concentrates on the case in which variables are CI(1,1) [Enders (1996), p.152]. The notion of cointegration can be equated with a linear equilibrium constraint between the integrated variables to be considered, leading to a class of models, known as error-correcting in the econometrics literature.

4.2.1a *Close Relationship between Cointegration and Linear Equilibrium Constraints*

Macroeconomic time series often appear to be I(1). Their behaviour may be similar to that of a (multivariate) random walk. A single equilibrium, or attractor, for an economic system to converge through time might be characterised as

$$\underline{\alpha}'\underline{Y}_t = 0, \tag{4.2}$$

so that the univariate quantity ε_t given by (4.1) can be interpreted as the extent to which the system is out of equilibrium at any point in time, and may thus be called the equilibrium error. Here, the term equilibrium can be thought of as a stable target towards which the economy, acting on countervailing forces, is aiming whilst in disequilibrium [Granger (1987), pp.5-6]. Provided that each component of $\underline{Y}_t$ is I(1), then if, and only if, the equilibrium error

ε_t is I(0) will the idea of a target equilibrium have any impact.[1]

Here the CV $\underline{\alpha}$ is not unique, since it could be multiplied by any nonzero scalar without affecting the equality in (4.2). However, it is the exception rather than the rule that there exists an $\underline{\alpha}$ which makes $\varepsilon_t \sim I(0)$, because economic equilibria do not always prevail, even if the presupposition of the same order of integration is satisfied.

In the more general case, with any values of K, d, b, for K>2, $\underline{\alpha}$ need not be unique, as there may exist more than one specific linear relationship amongst the variables. It then follows that if there are r linearly independent cointegrating vectors (CVs), then r can be, at most, equal to K-1, so that $r \leq K-1$. The cointegration combination stated above will then become

$$\underset{(r\times K)}{\underline{\alpha}'}\underset{(K\times 1)}{\underline{Y}_t} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1K} \\ \vdots & & \\ \alpha_{r1} & \cdots & \alpha_{rK} \end{bmatrix} \begin{bmatrix} Y_{1t} \\ \vdots \\ Y_{Kt} \end{bmatrix} = \underset{(r\times 1)}{\begin{bmatrix} \sum_{i=1}^{K} \alpha_{1i} Y_{it} \\ \vdots \\ \sum_{i=1}^{K} \alpha_{ri} Y_{it} \end{bmatrix}} = \underline{\varepsilon}_t, \qquad (4.3)$$

where $\underline{\alpha}$ is a (K×r) cointegrating or equilibrium matrix with its columns forming r distinct CVs in the r-dimensional cointegrating space and its rank r being referred to as the cointegrating rank of $\underline{Y}_t$; and

$\underline{\varepsilon}_t$ is a (K×1) (zero-mean) vector of stochastic variables integrated of order less than d.

For a given set of variables, we may be interested not only in a single CV, but in a set of CVs, specified by the (K×r) matrix $\underline{\alpha}$.

---

[1] It may be recalled that if ε_t is I(0) the expected time between zero crossings is finite, suggesting that in this case ε_t will rarely drift far from target $\varepsilon_t = 0$ and equilibrium will occasionally occur. Conversely, if there is an equilibrium, the $\underline{Y}$ variables must move closely together as a group over the long run. The selected interrelated variables will adjust to one another, by means of the CV $\underline{\alpha}$, such that the resulting stochastic error ε_t will be I(0).

This matrix is also not uniquely defined by the set of equations shown. The reason is that any elementary linear transformation of the column vectors involved (e.g. multiplying one column by a nonzero scalar, or adding to one column a nonzero scalar multiple of another column) will change its form or pattern, but not its rank, and the resulting $\underline{\varepsilon}_t$ will still be I(0). Such a matrix is said to be 'closed' with respect to addition and with respect to multiplication by a scalar. Moreover, although the loss of uniqueness adds extra flexibility, it brings some interpretational problems in terms of equilibria. To illustrate the problem, we will consider first a simple K=3 and r=2 case with the help of three-dimensional space analytic geometry and then, by intuition or analogy, extend this to the more general case.[2]

When a pair of independent CVs, $\underline{\alpha}_1$ and $\underline{\alpha}_2$, occur between three series, Y_1, Y_2 and Y_3, each of which is I(1), we shall write $\varepsilon_t(\underline{\alpha}_i) = \underline{\alpha}_i'\underline{Y}_t \sim I(0)$, i=1, 2, corresponding to the equilibrium relationships

$$\underline{\alpha}_1'\underline{Y}_t = \alpha_{11}Y_{1t} + \alpha_{12}Y_{2t} + \alpha_{13}Y_{3t} = 0, \qquad (4.4a)$$

and

$$\underline{\alpha}_2'\underline{Y}_t = \alpha_{21}Y_{1t} + \alpha_{22}Y_{2t} + \alpha_{23}Y_{3t} = 0. \qquad (4.4b)$$

These two linear equations will, in turn, correspond to two different planes $\mathcal{P}_1$ and $\mathcal{P}_2$ in 3-D space analytic geometry with a set of coordinate axes Y_1 Y_2 Y_3, which will both cut through the origin in the situation now being considered.[3] If each term of the equation in (4.4a) and (4.4b) is multiplied by arbitrary nonzero scalars λ_1 and λ_2, respectively, we obtain $\lambda_1\underline{\alpha}_1'\underline{Y}_t=0$ or $\lambda_2\underline{\alpha}_2'\underline{Y}_t=0$. These are equations of the same two planes $\mathcal{P}_1$ and $\mathcal{P}_2$ in the space, whatever the scalars are, implying that the cointegrating vectors are not unique. On the other hand, the equilibrium relationships existing among a set of cointegrated variables are unchanging. Further, any linear combination of pairwise independent $\underline{\alpha}_1$ and $\underline{\alpha}_2$ also yields a

---

[2] For this part, see Granger (1986b), pp.220-6, *inter alia*.

[3] A 3-vector of nonzero constants could be included in the equations (4.3), if needed, to make the mean of $\underline{\varepsilon}_t$ zero. This is equivalent to a parallel shift of the coordinate system from an old one $O(Y_1, Y_2, Y_3)$ to a new one $O'(Y_1', Y_2', Y_3')$ in 3-D space.

further cointegrating vector. It follows that

$$\begin{aligned}\varepsilon_t(\lambda) &= \lambda\varepsilon_t(\underline{\alpha}_1) + (1-\lambda)\varepsilon_t(\underline{\alpha}_2) \\ &= [\ \lambda\underline{\alpha}_1' + (1-\lambda)\underline{\alpha}_2'\ ]\underline{Y}_t\end{aligned}$$

will also be I(0), for any λ. Thus, the equilibrium relations are not uniquely identified, and the constructed planes will generate, or span, the whole three-space, indicating that the equilibrium error cannot be strictly interpreted as the deviations from a particular pair of equilibrium relationships. The only invariant, or stable attractor, is the intersection line $\mathcal{L}$: $\mathcal{P}_1 \wedge \mathcal{P}_2$ in the (Y_1, Y_2, Y_3) space defined by $\{\varepsilon_t(\underline{\alpha}_1)=0\} \cap \{\varepsilon_t(\underline{\alpha}_2)=0\}$ or $\{\varepsilon_t(\lambda_1)=0\} \cap \{\varepsilon_t(\lambda_2)=0\}$; $\forall\lambda_1,\lambda_2$; $\lambda_1 \neq \lambda_2$; refer to Fig.4.1.

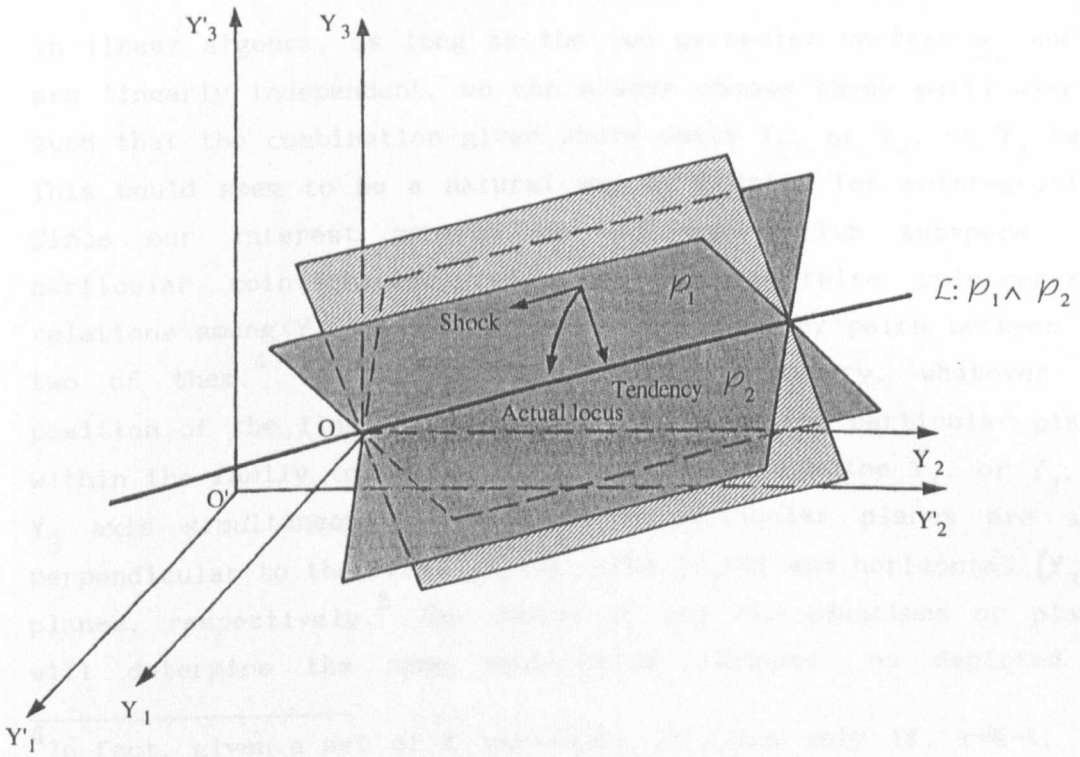

Fig.4.1 The determination of equilibrium sub-space in 3-D space.

This line, passing through the origin 0, is actually a particular region of the possible outcome space and will be called the equilibrium sub-space [see Granger (1986b), p.222]. In most time periods, the trivariate joint process $\underline{Y}_t$ whose components drift widely will not be on or close to the line $\mathcal{L}$, but might have a

generalised preference towards $\mathcal{L}$ — a certain part of the whole process space. The absolute value of equilibrium error might thus be expressed as the distance of the system from the equilibrium sub-space, and the process $\underline{Y}_t$ is said to be back to the sub-space or in a particular state of statistical equilibrium if, and only if, the multiseries $\underline{\varepsilon}_t$ are all zero [see Granger (1986b), p.226]. Accordingly, in order to represent adequately the long-run properties of a K-dimensional cointegrated process $\underline{Y}_t$ with cointegration rank r, $0<r<K$, it is important to find the right cointegration space spanned by the columns of $\underline{\alpha}$ or all possible distinct cointegration vectors. Any underestimate or overestimate of the true value of r (imposing too little or too much cointegration) may be misleading.

In linear algebra, as long as the two parameter vectors $\underline{\alpha}_1$ and $\underline{\alpha}_2$ are linearly independent, we can always choose three particular λs such that the combination given above omits Y_1, or Y_2, or Y_3 term. This would seem to be a natural way of testing for cointegration. Since our interest centres on the equilibrium subspace, not particular cointegration relations, the pairwise cointegration relations among Y_1, Y_2 and Y_3 can be replaced by pairs between any two of them.[4] In 3-coordinate analytic geometry, whatever the position of the line $\mathcal{L}$, we can always find three particular planes within the family to contain both the line $\mathcal{L}$ and the Y_1, or Y_2, or Y_3 axis simultaneously. These three particular planes are also perpendicular to the front $(Y_1=0)$, side $(Y_2=0)$ and horizontal $(Y_3=0)$ planes, respectively.[5] The choice of any two equations or planes will determine the same equilibrium subspace, as depicted in

---

[4] In fact, given a set of K variables, if, and only if, r=K-1, can all r (distinct) cointegration relations be transformed into the simplest r bivariate linear combinations between anyone of the K variables and each of the rest K-1 variables. Normally, the $r(\leq K-1)$ cointegration relations will each be characterised by a subset of, at most, $K-r+1(\geq 2)$ variables. In such linear transforms, no matter what the particular CVs are, the number of CVs (i.e. the dimension of the cointegration space) r would not change.

[5] Similarly, in the new coordinate system $0'(Y_1', Y_2', Y_3')$, the line $\mathcal{L}$ may not go through the origin 0', but the three particular planes would be parallel to the Y_1', or Y_2', or Y_3' axis.

Fig.4.2.

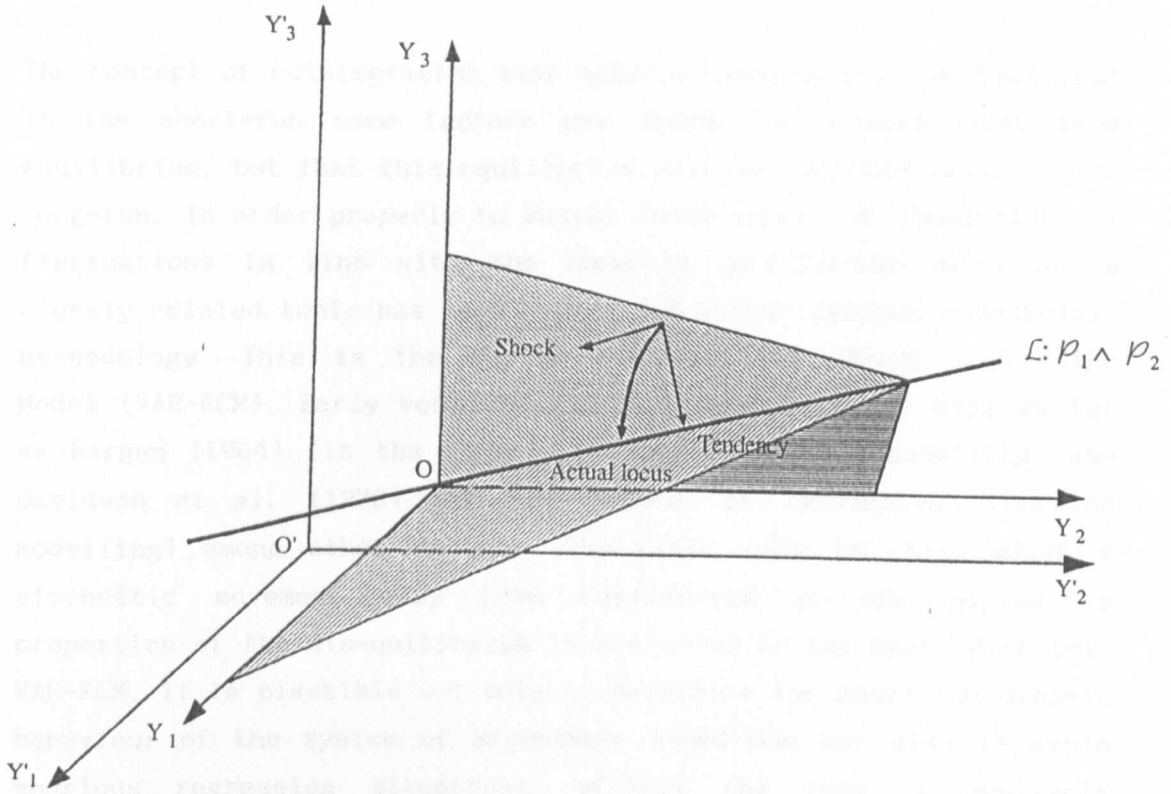

Fig.4.2 Three particular planes corresponding to simplified equations in 3-D space.

These ideas considered in three-space may be generalized so that for any K, r, the preferred equilibrium subspace will be a hyper-plane of dimension K-r(≥1) [see Granger (1986b), p.222]. In applications, the cointegrating matrix (or the matrix of cointegrating vectors) $\underline{\alpha}$ and its rank r are unknown. Whether or not cointegration occurs, how many distinct CVs there are and what the scaling matrix or vector will be are empirical questions and are not necessarily determined by economic theory. In particular, the higher the K, the more extensive is the number of possible combinations and the more difficult the testing will be.

The importance of the empirical existence of cointegration, therefore, stems from the fact that tests for cointegration are tests for suspected long-run equilibrium relationships among the economic variables of interest. Hence, cointegration should be considered a necessary condition for the long-run components of

integrated variables to obey equilibrium constraints while the short-run components capture flexible dynamics.[6]

The concept of cointegration also makes allowance for the fact that in the short-run some factors may shock the economy away from equilibrium, but that this equilibrium will be restored again in the long-run. In order properly to adjust these short-run disequilibrium fluctuations in line with the long-run equilibrium solution, a closely related topic has been propounded within dynamic econometric methodology. This is the Vector Autoregressive Error Correction Model (VAR-ECM). Early versions can be traced back at least as far as Sargan (1964) (in the context of wage and price modelling) and Davidson *et al.* (1978) (in the context of consumption function modelling) among other writers. The basic idea is that, given a stochastic movement away from equilibrium in one period, a proportion of the disequilibrium is corrected in the next. With this VAR-ECM, it is plausible not only to determine the short-run dynamic behaviour of the system of adjustment equations but also to avoid spurious regression situations, without the loss of desirable long-run information.

4.2.1b *Close Relationship between Cointegration and Vector Autoregressive Error Correction Models*

We now demonstrate how cointegration can arise from a finite order VAR representation, commonly used in econometrics as a convenient approximation to a system of infinite order. Consider a K-dimensional Gaussian VAR(p), as in (2.3), with fixed initial values $\{\underline{Y}_{1-p}, \ldots, \underline{Y}_0\}$

$$\underline{Y}_t = \underline{C} + \sum_{i=1}^{p} \underline{A}_i \underline{Y}_{t-i} + \underline{u}_t, \qquad t = 1, \ldots, T, \qquad (4.5)$$

[6]Strictly speaking, a necessary and sufficient condition for equilibria to hold is $\underline{\varepsilon}_t \sim I(0)$, not cointegration; the latter is only a necessary one for the process to be in such a particular state, as has been pointed out by Granger (1986b), Engle & Granger (1987) and others. This is because the reduced-rank restrictions implied by cointegration cannot guarantee that the resultant cointegrating linear combinations $\underline{\alpha}'\underline{Y}_t = \underline{\varepsilon}_t$ are I(0) unless the original series are known to be I(1).

where all the symbols have their usual meanings. Although most of the time the simplifying assumptions, such as normality, independence, and homoscedasticity, are not crucial, the assumption that $\{\underline{C},\underline{A}_t,\ldots,\underline{A}_p,\Sigma_u\}$ are time-invariant parameters is fundamental. This assumption will be investigated in the empirical applications in subsequent chapters.

Recall that a necessary and sufficient condition for the process to be stationary is that all characteristic roots of the polynomial equation $\det(\underline{A}_p(\xi))=0$ are outside the unit circle. Thus, if the process is nonstationary there must be at least one root lying on or inside the unit circle. Due to the presence of a single unit root in most integrated macro series, we exclude explosive variables by assuming that none of the roots lie inside the unit circle and also rule out $\underline{Y}_t$ being integrated of order 2, though the latter could be handled, should the need arise.

Now suppose the process encountered here is nonstationary with

$$\begin{aligned}|\underline{A}_p(\xi)| &= |I_K - \sum_{i=1}^{p} \underline{A}_i \xi^i| \\ &= \prod_{i=1}^{Kp} (1 - \lambda_i \xi) \\ &= 0, \qquad \text{for } |\xi| = 1.\end{aligned}$$

Since $\{\lambda_i\}$, $i=1,\ldots,Kp$, stand for the reciprocals of the roots of the characteristic equation, one or more of them must be equal to 1 in absolute value. A multivariate time series model therefore needs to be constructed in such a way that both valid reduced-rank restrictions and unit roots are imposed explicitly on its various components in order for the joint modelling to reflect common features such as seasonals and cycles which are felt to be present in the aggregate data.

Since $|\underline{A}_p(1)|=|I_K-\sum_{i=1}^{p}\underline{A}_i|=0$, the $K\times K$ long-run multiplier matrix

$$\underline{A}_p(1) = I_K - \sum_{i=1}^{p} \underline{A}_i \tag{4.6}$$

is singular and has a reduced rank. Suppose $rk[\underline{A}_p(1)]=r$, $0<r<K$, so

that the impact matrix $\underline{A}_p(1)$ can be factorised into the product of two suitable matrices $\underline{\omega}$, $\underline{\alpha}'$ with both $\underline{\omega}$ and $\underline{\alpha}$ being of dimension K×r and of full column rank. Three cases arise from the study of cointegration constraints.

Case (a): r=0 — a borderline case, in which $\underline{A}_p(1)$ is a null matrix with K unit roots imposed. All K variables in $\underline{Y}_t$ are I(1), and a stationary unrestricted VAR(p-1) model represented purely in first differences is acceptable.

Case (b): r=K — a stationary case, in which $\underline{A}_p(1)$ is a full-rank matrix with no unit roots imposed. All K variables in $\underline{Y}_t$ are I(0), and a stationary unrestricted VAR(p) process expressed entirely in levels is admissible.

Case (c): 0<r<K — an intermediate case, in which $\underline{A}_p(1)$ is a singular matrix with K-r unit roots imposed. There are, at most, r cointegrated linear combinations of the elements of $\underline{Y}_t$, and a stationary restricted VAR-ECM is appropriate.

In general, if $\underline{Y}_t$ consists of cointegrated I(1) variables, the stationary VAR-ECM formulation with Gaussian innovations can be achieved by subtracting $\underline{Y}_{t-1}$ from both sides of (4.5) so that the impact matrix enters explicitly

$$\begin{aligned}
\nabla\underline{Y}_t &= \underline{Y}_t - \underline{Y}_{t-1} \\
&= \underline{C} - (I_K - \underline{A}_1 - \ \cdots \ - \underline{A}_p)\underline{Y}_{t-1} \\
&\qquad - (\underline{A}_2 + \ldots + \underline{A}_p)(\underline{Y}_{t-1} - \underline{Y}_{t-2}) \\
&\qquad\qquad \cdots \\
&\qquad\qquad - \underline{A}_p(\underline{Y}_{t-p+1} - \underline{Y}_{t-p}) + \underline{u}_t \\
&= \underline{C} - (I_K - \sum_{i=1}^{p} \underline{A}_i)\underline{Y}_{t-1} - (\sum_{i=2}^{p} \underline{A}_i)\nabla\underline{Y}_{t-1} - \ldots - \underline{A}_p\nabla\underline{Y}_{t-p+1} + \underline{u}_t, \\
&\qquad t = 1, \ldots, T,
\end{aligned}$$

where ∇ is the first-difference operator (1-L). Using (4.6) and defining

$$\underline{A}_j^* = -\sum_{i=j+1}^{p} \underline{A}_i, \qquad j = 1, \ldots, p-1,$$

the VAR-ECM can be written as

$$\begin{aligned}\nabla \underline{Y}_t &= \underline{C} - \underline{A}_p(1)\underline{Y}_{t-1} + \underline{A}_1^*\nabla\underline{Y}_{t-1} + \ldots + \underline{A}_{p-1}^*\nabla\underline{Y}_{t-p+1} + \underline{u}_t \\ &= \underline{C} - \underline{A}_p(1)\underline{Y}_{t-1} + \sum_{j=1}^{p-1}\underline{A}_j^*\nabla\underline{Y}_{t-j} + \underline{u}_t, \quad t = 1, \ldots, T. \qquad (4.7)\end{aligned}$$

Also, the characteristic equation of the model (4.7) can be rewritten as

$$\begin{aligned}|\underline{A}_p(\xi)| &= |(1-\xi)I_K + \underline{A}_p(1)\xi - (1-\xi)\sum_{j=1}^{p-1}\underline{A}_j^*\xi^j| \\ &= 0, \qquad \text{for } |\xi| = 1.\end{aligned}$$

This implies that a transformation to I(0) space can be achieved in terms of differences and cointegrating linear combinations of the components of $\underline{Y}_t$.

Within this framework the long-run impact matrix $\underline{A}_p(1)$ has been decomposed as $\underline{A}_p(1)=\underline{\omega\alpha}'$. The matrix $\underline{\alpha}$ is, as before, the cointegrating matrix comprising r distinct column cointegrating vectors. $\underline{\omega}$ is the loading matrix containing, for all K equations in the system, the adjustment (or error-correction) parameters attached to each CV. This shows how fast the variables change in response to a disequilibrium.[7] See, for example, Clements and Mizon (1991, pp.895-6), and Urbain (1995, pp.181-3). Thus, (4.7) becomes

$$\begin{aligned}\nabla \underline{Y}_t &= \underline{C} - \underline{\omega\alpha}'\underline{Y}_{t-1} + \sum_{j=1}^{p-1}\underline{A}_j^*\nabla\underline{Y}_{t-j} + \underline{u}_t \\ &= \underline{C} - \underline{\omega\varepsilon}_{t-1} + \sum_{j=1}^{p-1}\underline{A}_j^*\nabla\underline{Y}_{t-j} + \underline{u}_t, \quad t = 1, \ldots, T, \qquad (4.8)\end{aligned}$$

or

$$\underline{\omega\varepsilon}_{t-1} = \underline{C} - \nabla\underline{Y}_t + \sum_{j=1}^{p-1}\underline{A}_j^*\nabla\underline{Y}_{t-j} + \underline{u}_t, \quad t = 1, \ldots, T.$$

All terms in the right-hand side involving the difference operator ∇

[7] Here again, the two matrices $\underline{\alpha}$ and $\underline{\omega}$ are not unique since, for any nonsingular r×r matrix $\underline{F}$, we can define $\underline{\alpha}^*=\underline{\alpha F}'$ and $\underline{\omega}^*=\underline{\omega F}^{-1}$ so that $\underline{\omega}^*\underline{\alpha}^{*\prime}=\underline{\omega\alpha}'$. However, it will be shown that closed form expressions for the ML estimators of $\underline{\alpha}$ and $\underline{\omega}$ are available.

are stationary, and the error term $\underline{u}_t$ is stationary by assumption. This implies that the error-correction term $\underline{\omega}\underline{\varepsilon}_{t-1}$, being a linear combination of the stationary terms only, must also be stationary and it remains stationary upon pre-multiplication by $(\underline{\omega}'\underline{\omega})^{-1}\underline{\omega}'$. Hence, if $\underline{\omega}\neq 0$, the 'equilibrium error' $\underline{\varepsilon}_t=\underline{\alpha}'\underline{Y}_t$ is stationary and thus each row of $\underline{\alpha}'\underline{Y}_t$ represents a cointegrating relation. It has been proved in Granger's Representation Theorem (Granger, 1983) that not only must integrated data generated by an error-correction model be cointegrated but the converse is also true. That is, if the components of $\underline{Y}_t$ are cointegrated of order (1,1), then there must exist a generating mechanism having the error-correction form.[8]

Similarly, an alternative yet isomorphic error correction version for the process (4.5) can be derived in the fashion of Johansen (1988, 1991), and Johansen and Juselius (1990, 1992), i.e.

$$\begin{aligned}\nabla\underline{Y}_t &= \underline{C} + \sum_{j=1}^{p-1}\underline{A}_j^{\dagger}\nabla\underline{Y}_{t-j} - \underline{A}_p(1)\underline{Y}_{t-p} + \underline{u}_t \\ &= \underline{C} + \sum_{j=1}^{p-1}\underline{A}_j^{\dagger}\nabla\underline{Y}_{t-j} - \underline{\omega}\underline{\varepsilon}_{t-p} + \underline{u}_t, \qquad t = 1, \ldots, T, \end{aligned} \tag{4.9}$$

where

$$\underline{A}_j^{\dagger} = -I_K + \sum_{i=1}^{j}\underline{A}_i, \qquad j = 1, \ldots, p-1.$$

In general, by rearranging terms, any set of lags of the $\underline{\varepsilon}$ can be reproduced in a like manner by

$$\begin{aligned}\nabla\underline{Y}_t &= \underline{C} + \sum_{j=1}^{l-1}\underline{A}_j^{\dagger}\nabla\underline{Y}_{t-j} - \underline{A}_p(1)\underline{Y}_{t-1} + \sum_{j=1}^{p-1}\underline{A}_j^{*}\nabla\underline{Y}_{t-j} + \underline{u}_t \\ &= \underline{C} + \sum_{j=1}^{l-1}\underline{A}_j^{\dagger}\nabla\underline{Y}_{t-j} - \underline{\omega}\underline{\varepsilon}_{t-1} + \sum_{j=1}^{p-1}\underline{A}_j^{*}\nabla\underline{Y}_{t-j} + \underline{u}_t, \\ & \qquad l = 1, \ldots, p, \qquad t = 1, \ldots, T, \end{aligned} \tag{4.10}$$

where

$$\underline{A}_j^{\dagger} = -I_K + \sum_{i=1}^{j}\underline{A}_i, \qquad j = 1, \ldots, l-1,$$

[8] In theory, the proposition presented here can be generalized to any values of d and b. For a proof of the theorem see, *inter alia*, Engle & Granger (1987), Hylleberg & Mizon (1989) and Johansen (1988, 1991).

and

$$\underline{A}_j^* = -\sum_{i=j+1}^{p} \underline{A}_i, \qquad j = 1, \ldots, p-1.$$

Particularly, when l=1, $\underline{A}_j^{\dagger}|_{j=0}$ is undefined and taken as zero, Eq.(4.10) then reduces to (4.8). Likewise, when l=p, $\underline{A}_j^*|_{j=p}$ is undefined and treated as zero, Eq.(4.10) then simplifies to (4.9). Therefore, any type of gradual partial adjustment of the short-run dynamics toward a long-run equilibrium can be allowed via the introduction of this further possible exogenous factor, the equilibrium error, that arises from the concept of cointegration.

When a set of variables are cointegrated, their changes may depend not only upon the past values of the variables in difference terms, if present, but also upon the past values of the equilibrium errors in level terms, if $\underline{\omega} \neq 0$. The change in the jth component of $\underline{Y}_t$ may be better expressed as a weighted average of lagged changes in all components less r lagged error correction terms (the jth row vector of $\underline{\omega}$ times the lagged $\underline{\varepsilon}_t$) plus the jth component of $\underline{u}_t$ without moving average components. The term EC, therefore, comes from the fact that if the economy tends to seek equilibrium, then it is expected that the variables involved should be influenced by the extent to which the economy is out of equilibrium. In an ECM, some fixed proportion, $\underline{\omega}$, of the disequilibrium, $\underline{\varepsilon}_{t-l}$ (l=1,...,p), in period t-1 would be corrected in the current period t in aiming at the long-run relationship or the equilibrium subspace defined previously. The relevant coefficients of the EC term would appear to be negatively correlated. See Lütkepohl (1991a), Chapter 11.

In such a specification, a long-run stationary relationship holding between variables will necessarily make $\underline{A}_p(1)$ nonzero (r>0). In this case, a pure VAR in the differences of the data will be inefficient, as it omits important long-run information [Doan (1996), p.8-3]. However, the long-run cointegrating restrictions will also make $\underline{A}_p(1)$ singular (r<K). In this case, a simple VAR in the levels of the series will be inadequate from an innovation accounting point of view, as it ignores crucial cross-equation parametric constraints.

To quote Engle and Granger (1987, p.259):

> *...vector autoregressions estimated with cointegrated data will be misspecified if the data are differenced, and will have omitted important constraints if the data are used in levels. Of course, these constraints will be satisfied asymptotically but efficiency gains and improved multistep forecasts may be achieved by imposing them.*

In essence, it is this feature of cointegrating restrictions that links it with the analysis of spurious regressions. This is because in the system of equations (4.9), every term (either the lagged levels implied by the presence of $\underline{\varepsilon}_{t-p}$ or the lagged differences reflected by the appearance of $\nabla\underline{Y}_t$) is I(0) when cointegration exists. Hence, conventional statistical inference can be carried out and spurious regressions of the Granger-Newbold type can be avoided. Whether or not the residual in a regression between integrated aggregates is empirical white noise, or at least stationary, is thus a matter of importance [see Mills (1993), Chapter 6, especially pp.170-2]. Moreover, a test for cointegration leading to this stationarity is, to all intents and purposes, a pre-test to get around the alternative hypothesis of spurious level combinations.

In contrast with previous studies, although the model looks at first sight similar in form to a traditional unrestricted VAR in differences, it is distinct in kind. One major difference is that the VAR-ECM is an open multivariate system rather than a closed one because the equilibrium error is included as an exogenous variable, along with the constant term, to let in potential information contained in the long-run solution of the process. Another is that the VAR-ECM is effectively nonlinear rather than linear in $\underline{\alpha}$ because $\underline{\omega}$ is unknown in the composite long-run parameter matrix and must therefore be estimated simultaneously with 'unscrambled' $\underline{\alpha}$ as well as with the other autoregressive coefficients.

4.2.2 A Further Useful Modification of the Model with the Possibility of Both Deterministic and Stochastic Nonstationarities

We now consider augmentation of the basic model to incorporate deterministic features via

$$\nabla \underline{Y}_t = \underline{C} + \underline{\Psi} \underline{D}_t + \sum_{j=1}^{p-1} \underline{A}_j^{\dagger} \nabla \underline{Y}_{t-j} - \underline{A}_p(1) \underline{Y}_{t-P} + \underline{u}_t, \qquad (4.11)$$
$$t = 1, \ldots, T.$$

Here $\underline{D}_t$ contains, except for a constant, any deterministic components [e.g., a linear trend (when $\underline{\Psi D}_t = \underline{\delta} t$) or a quadratic trend, a regime shift, a seasonal, or other dummies] that are exogenous to the VAR system under consideration, to ensure that the disturbances $\underline{u}_t$ are as close to being white noise as possible. We note that even if the nonstationary influence of deterministic trends is straightforwardly allowed for, allowance should still be made for the possibility of stochastic trends in the internal dynamics; i.e. some roots in the characteristic polynomial lie in the neighbourhood of unity. Indeed, the more comprehensive is the model, the greater is the internal consistency with which an introduction of exogenous changes in level or trend may strengthen model-based predictions across shifts in policy regimes [Wallis (1989), p.32].

The combined version in (4.11) with $\underline{\Psi D}_t = \underline{\delta} t$ incorporates two competing sources of trend: a deterministic trend in mean; and a stochastic trend in covariances. Six main specializations of (4.11) can be categorized in terms of the trend specification imposed on the data, as given in Table 4.1.

Table 4.1: Classification of various cases involved in the common model (4.16)

$\underline{\delta}$ \ r^*	r=K	r=0	0<r<K
$\underline{\delta}=0$	A stationary VAR process	A difference-stationary (DS) VAR process	A stationary VAR-ECM process with no trend in mean
$\underline{\delta}\neq 0$	A trend-stationary (TS) VAR process	A difference- and trend-stationary (DTS) VAR process	A stationary VAR-ECM process about a trend in mean

*$r = rk[\underline{A}_p(1)]$ is the number of cointegrated linear combinations.

The case for including components, such as trends or shifts in the mean or in the covariances or specific seasonals, by polynomial regression, will largely depend on an initial analysis of both the normalized and non-normalized variables in each equation. The

analysis of the former may indicate which of the movements in the normalized variable are accounted for by the non-normalized variables, while the analysis of the latter may indicate what properties a potential non-normalized variable has in common with the variable being forecast. A minimum condition needed for an acceptable model is that a simulation of the right-hand side should be able to reflect the salient characteristics appearing in the left hand side variable. The inclusion of a dummy or seasonal explanatory variable therefore opens up the possibility that trend or seasonal variations in the normalized variable are capable of being explained, along with other kinds of variations, by the variations in the regressors so obtained [cf. Harvey (1992), p.390].

In reality, the danger of over-sophisticating a nonstationary stochastic process may be avoided by the addition of a trend component to all equations and letting the data determine whether or not it is needed. However, in so doing, care has to be exercised in evaluating the sensitivity of inference based on VARs with potentially different trend assumptions, as the discriminatory power of the tests used in identifying trends in economic time series is limited. The modification of the critical values in joint hypothesis testing where conventional asymptotics apply may be necessary or inevitable in many empirical investigations. See, for example, Ohanian (1988, 1990).

§4.3 Error-Learning, Parameter Constancy and the Lucas Critique

Usually, the forecasts made by a statistically well-specified VAR-ECM formulation on the basis of the available information are adaptive, or error-learning. By adaptive we mean that the prediction of the future value of variables is revised only mechanically by some fixed proportion of the extent to which expectations are not realised in the current period. Such an adjustment created in response to systematic past errors can improve forecast accuracy, but also reflects the problem that the model has not extracted all the useful information in the historical data, leaving scope for

further development.

The hypothesis by which predictions are made without systematic mistakes is that of rational expectations formed on the basis of all information available to agents about future events. The rational expectations hypothesis implies a number of testable restrictions on the parameters, but tests of such (overidentifying) restrictions are actually difficult to operate. As Artis (1988, pp.5-6) has stated:

> *... testing the rationality of the forecasts involves assessing the contributions of 'innovations' in the exogenous variables. ... it is less easy to do this for fiscal policy and appears not to be feasible for monetary policy.*

While accepting the expectations variant of the Lucas (1976a) critique, Sims challenges rational expectations econometrics even further by casting doubt on the validity of supposed *a priori* knowledge in achieving acceptable identifying restrictions in structural models. He writes (1980a, p.7):

> *It is my view, however, that rational expectations is more deeply subversive of identification than has yet been recognised.*

He continues:

> *In the presence of expectations, it turns out that the crutch of a priori knowledge of leg lengths is indispensable, even when we have distinct strictly exogenous variables shifting supply and demand schedules.*

Indeed, the in-built assumption of expectations implies that agents know the model, but this may be untrue for many agents. It would therefore be more pragmatic to find a compromise operational method between the merely error-learning and the fully rational. The explicitly Bayesian vector autoregressive (BVAR) forecasting procedure is one way of meeting this requirement.

Moreover, the key assumption underlying the VAR-ECM technique is, as mentioned earlier, parameter constancy at the cost of mechanically introducing error-correction feedbacks and, possibly, dummy variables to equations. Nevertheless, this is still a strong assumption that may not be acceptable in the light of the Lucas critique. The Lucas critique from the perspective of rational

expectations with competitive equilibrium models, in which all markets are assumed to clear immediately, provides a reminder of the inherent limitations of macroeconomic models, whose behavioural parameters would not remain invariant, but would be allowed to vary over time. In addition, not only are the values of all the coefficients in the system subject to the influence of alternative government policies but so are the interdependencies and degree of instability of these coefficients. Hence, parameter non-constancy is a problem, not easily overcome, though the likely force of the Lucas critique is essentially an empirical question.

One viable method of handling parameter non-constancies in models where there are stochastic trends is the Bayesian approach to inference. Bayesian analysis has two advantages. One is that it can formally integrate order selection and hypothesis testing in a cointegrated Gaussian VAR to produce a joint coherent model selection (or model discrimination) and prediction procedure. The other is that, if properly formulated, it can represent uncertainty about estimation results in ways not available with the classical approach. As such, it may yield good performance in VARs with some unit roots and some cointegration.

Furthermore, the existence of r independent cointegrating relations reduces the number of parameters in the impact matrix, $\underset{\sim}{A}_p(1)$, from K^2 to $Kr+(K-r)r$, or (the number) of parameters in the multivariate cointegration system, (3.11), from $K+K^2p$ to $K+Kr+(K-r)r+K^2(p-1)$.[9] However, VAR-ECMs, though behaving quite differently from UVARs, have not fundamentally overcome the central forecasting issue of overparameterization that the larger the model, the more the

[9]The expression means that among the K rows/columns of the $\underset{\sim}{A}_p(1)$ matrix, only r rows/columns are linearly independent and the rest (K-r) rows/columns are each certain linear combination of these r independent rows/columns. Consequently, the space spanned by the columns of $\underset{\sim}{\alpha}$ (termed the cointegration space $sp(\underset{\sim}{\alpha})$) is the row space of $\underset{\sim}{A}_p(1)$, and the space spanned by the columns of $\underset{\sim}{\omega}$ (called the adjustment space $sp(\underset{\sim}{\omega})$) is the column space of $\underset{\sim}{A}_p(1)$. It also means that $K+Kr+(K-r)r+K^2(p-1)-[K+K^2p]=-(K-r)^2<0$, for $0<r<K$. Accordingly, the imposition of the cointegrating restrictions on the parameters of the model could save on the number of parameters to be estimated.

insignificant parameters have to be estimated, and thus the poorer estimates and forecasts will result. This problem can be addressed by the use of Bayesian methods with a proper informative prior specified for the VAR parameters. Apart from economizing on the dimension of the parameter space, BVARs are also likely to deliver better multi-period forecasts associated with the best setting of the hyperparameters.

§4.4 Conclusion

This chapter explores the estimation of a VAR containing nonstationary variables. At one time, it was believed that stochastic trends of nonstationary variables used in a regression analysis could be removed by differencing. The resulting stationary series could be estimated using traditional techniques. However, the conventional wisdom of differencing all nonstationary variables in a multivariate context was inappropriate. Recent development of time-series analysis reveals that there is much potential, in a statistical sense, for taking into account the hidden common stochastic trends, i.e. cointegration. In this context, studies of cointegrated multivariate time series have proved very useful in system modelling and forecasting [for instance, Boswijk (1995), Engle *et al.* (1993), Ericsson (1995), Gali (1992), Wickens (1996), Söderlind & Vredin (1996), and Urbain (1993, 1995)].

When the considered data series are small and nonstationary, another important undertaking is the development of a formal and flexible BVAR model embodying stochastic prior coefficient restrictions. Nonetheless, there seems to be no one generally acceptable path to a good model. The available BVAR and EC alternatives each have their supporters and critics, and are subject to controversies [Granger (ed) (1991), Chapter 1]. It may be that the two models have complementary, rather than competing roles. Hence, they might be combined to form a more general BVAR-EC model, thereby improving forecast accuracy. This is potentially of practical importance, and will be considered in the following chapters.

CHAPTER 5

NONSTATIONARY BAYESIAN VECTOR AUTOREGRESSIONS (BVARs) WITH BOTH PRIOR AND COINTEGRATION RESTRICTIONS

§5.1 Introduction

An attractive feature of the Bayesian probabilistic approach is that, apart from taking parameter uncertainty into account, it provides us with a coherent procedure in which prior information can be formally expressed and blended with sample data. The passage, via Bayes' theorem, from the combination of prior and sample information to the posterior probability density function (pdf) is itself part of a learning and discovering process. The prior distribution, being a component of posterior pdf, will be assumed a multivariate normal for the coefficients of the vector autoregression. The Gaussian likelihood, being another component, will be similar in form to the posterior odds when the sample is highly informative about the coefficients. In a linear VAR with normal disturbances, the Gaussian likelihood, holding fixed the covariance matrix of the disturbances and the initial observations, is proportional to a normal pdf. Hence, the posterior distribution, which is just the likelihood function weighted by the prior pdf, will be asymptotically normally distributed even when the true process has a unit root. This is especially useful in our practical empirical research, as it allows us to make exact inference, conditional on initial observations, relatively easily. Moreover, this procedure can be construed as formalizing a coherent methodological framework for forecasters to document, discuss and exchange their beliefs and, thereby, can help make macroeconomic forecasting more of a science and less of an art.

In this chapter, specification of finite BVAR models with both prior and cointegration restrictions is provided. Estimation of such models, especially for forecasting with the help of the Kalman Filter estimation method, is also described. The final section is

devoted to multi-step model-based forecasts. These considerations are applied in the forthcoming chapters.

§5.2 Specification of a Finite 'Open' BVAR Model

For simplicity, we confine ourselves here to the determination of prior and cointegration restrictions in a Bayesian specification framework. It is presumed that a pre-test of the number of lags and the number of variables has been done in an initial specification search as before. To this end, a sequential testing procedure, consisting of four steps, is carried out as follows.

(a) As a prerequisite, the tests for the order of integration or the number of unit roots of all the series appearing in the model must first be implemented singly with the aid of the augmented Dickey-Fuller (ADF) regression.

(b) For the given orders, the tests for the presence and number of cointegration relationships can then be executed on the basis of two popular test strategies. One is the residual-based ADF method proposed by Engle and Granger (1987) and the other is the system-based Full Information Maximum Likelihood (FIML) approach advanced by Johansen (1988, 1991). See also Johansen and Juselius (1990, 1992, 1994). These have asymptotically equivalent properties under certain conditions.

(c) For a given value of r (the dimension of the cointegration space), the estimates of the parameters of the CVs, used in estimating a multivariate dynamic model, can be performed via the use of the same two procedures.

(d) Two alternative types of joint prior structures in the style of Litterman can be constructed, compared and contrasted for a closed BVAR model in levels (with all the variables assumed to be jointly endogenous), and an open BVAR-EC model (with exogenous error-correction feedbacks).

Once cointegrating and prior restrictions have been determined sequentially for model specification, it is possible to estimate the model with the former placed on the long-run matrix, and the latter on the other parameters of the system discussed in section 5.3.

5.2.1 Determination of Cointegration Restrictions

In the econometrics literature, it is not assumed that cointegrating relationships are known *a priori*. Rather, the testing of the hypothesis of cointegration consists of two parts: tests for I(1) of the individual series; and if this is satisfied, testing for I(0) of a linear combination. These will be considered shortly.

5.2.1a *Testing for the Order of Integration*

The first step in testing for cointegration amongst the variables of interest is to determine the order of integration of the individual time series. Various statistical tests of unit roots are now available, but considering that many (perhaps most) macroeconomic data appear to be integrated of order one, we will concentrate explicitly on a simple, scalar version of the OLS regression test for a single unit root.

The simplest test for the hypothesis of a unit root in each component series, say Y_{jt}, $j=1,\ldots,K$, is the test that a regression of Y_{jt} on $Y_{j,t-1}$ yields a coefficient of unity. Such tests, based mainly on variants of an AR(1) model (with serially uncorrelated error terms), have been pioneered and modified by Fuller (1976) and Dickey and Fuller (1979, 1981) to an AR(m) scheme that is asymptotically valid in the presence of serial correlation. The generalised specifications with either an intercept or an intercept and a time trend can be given by

$$\nabla Y_{jt} = \rho_{j0} + \rho_{j1} Y_{j,t-1} + \sum_{i=1}^{m} \gamma_{ji} \nabla Y_{j,t-i} + \zeta_{jt} \qquad (5.1a)$$

and

$$\nabla Y_{jt} = \rho_{j0} + \rho_{j1} Y_{j,t-1} + \rho_{j2} t + \sum_{i=1}^{m} \gamma_{ji} \nabla Y_{j,t-i} + \zeta_{jt}, \qquad (5.1b)$$

respectively. These two modified regressions are often called augmented Dickey-Fuller (ADF) tests, which would have exactly the same asymptotic distributions as the ordinary Dickey-Fuller (DF) statistics, if the autocorrelations in ∇Y_{jt} were fully accounted for by the specified m-lag univariate AR processes. Three points should be made concerning the models (5.1a) and (5.1b) that are used to construct the ADF test statistics.[1]

(1) The sensitivity of the asymptotic distributions

As regards statistical inference, one might expect that a unit root test could be accomplished simply by using the classical t-ratio attached to $\hat{\rho}_{j1}$ on $Y_{j,t-1}$ to investigate the significance of $Y_{j,t-1}$ after OLS estimation of either (5.1a) or (5.1b), since $\rho_{j1}=0$ under the null hypothesis of a unit root.[2] However, this test is complicated when the data are actually generated by a nonstationary process. This is because the conventional asymptotic tests cannot be applied to some or all of the parameter estimates, even in large samples. The OLS estimate $\hat{\rho}_{j1}$ must be normalized by a factor T instead of $T^{1/2}$ and the corresponding cumulative distributions for T-norming of $\hat{\rho}_{j1}$ must be simulated numerically. Moreover, it turns out that the limiting distributions of the test statistics under the null hypothesis are especially sensitive to the values of nuisance parameters in both generalisations and, therefore, must be considered separately if a valid test of the unit root hypothesis is to be carried out.

To clarify the effects of differing nuisance parameters on the asymptotic behaviour of these two important generalisations, consider ADF regression (5.1a) without a time trend first. Two cases are commonly encountered under the unit root hypothesis.

---

[1] For a more comprehensive study of univariate unit root testing see, among others, Davidson & Mackinnon (1993), Chapter 20, and Godfrey (ed) (1992), Chapter 1.

[2] It is important to bear in mind that, in most cases, the published critical values for these statistics are lower-tail ones, since the alternative of interest is almost always that the process is stationary, not that it is explosive.

Case (a): if $\rho_{j0}=0$, i.e. there is no drift in the process, the appropriate tables of the adjusted critical values that permit the application of t tests to the coefficient on $Y_{j,t-1}$ are characterised by Monte Carlo simulations either in Fuller (1976, Table 8.5.2) or in Mackinnon (1990, Table 1).

Case (b): if $\rho_{j0}\neq 0$, i.e. there is a nonzero drift in the process, t-ratio regression test for $\rho_{j1}=0$ appears to converge to the standard normal N(0,1) distribution asymptotically [see Pagan & Wickens (1989), pp.968-9].

Now consider such effects in the second, ADF regression (5.1b) with a time trend. In this situation, there are also two cases of special interest under the null hypothesis being tested:

Case (a): if $\rho_{j2}=0$, whatever the values of ρ_{j0}, the asymptotic distributions of $t(\rho_{j1})$ and $F(\rho_{j1},\rho_{j2})$ are invariant over all values of ρ_{j0}. Reference to the DF distributions for t and F statistics, reported only for $\rho_{j0}=0$, should be made to achieve a correct test for the unit root null [see Fuller (1976) and Dickey & Fuller (1981)]. Whether ρ_{j0} is zero or nonzero will only affect the limiting distribution of ρ_{j2}: when $\rho_{j0}=0$, we use Table III in Dickey and Fuller (1981), whereas, when $\rho_{j0}\neq 0$, we use the top panel of Table 4 in Nankervis and Savin (1987).

Case (b): if $\rho_{j2}\neq 0$, whatever the values of ρ_{j0}, the conventional t and F statistics are found to be asymptotically valid, i.e. they will have the usual asymptotic normal and chi-squared distributions, respectively. See Nankervis and Savin (1987), the bottom panel of Table 1.

The fact that the distribution of $t(\rho_{j1})$ under the null is nonnormal in (5.1a) when $\rho_{j0}=0$ or in (5.1b) when $\rho_{j2}=0$ is primarily due to the mean, which is significantly negative. As ρ_{j0} and/or ρ_{j2} increases in absolute value or, equivalently, as σ_ζ goes to zero, the mean of the distribution shifts towards zero and the shape of the distribution becomes approximately symmetric. See Davidson and Mackinnon (1993), and Nankervis and Savin (1987), for details.

(2) The determination of the lag length

The main reason of expending the regressor set by including m additional lagged first differences is to handle serial correlation of unknown form in the errors. However, including too many lags reduces the power of the test to reject the null of a unit root, since the increased number of lags necessitates the estimation of additional parameters and a loss of degrees of freedom. On the other hand, too few lags will not appropriately capture the actual error process, so that ρ_{j1} and its standard error will not be well estimated [see Enders (1996), p.90].

In general, the value of desired truncation lag m should grow with T at a rate proportional to $T^{1/3}$ that allows for an adequate AR approximation of general autocorrelation. In practice, one approach is to start with a relatively long lag length and pare down the model by the usual significance tests, provided the true lag length is covered in the initial choice. A practical rate for selecting values of m is to ensure that the regression residuals $\hat{\xi}_{jt}$, j=1,...,K, are empirical white noise.

(3) The inclusion of a time trend

The inclusion of a linear time trend in regression models like (5.1b) is important because trend stationarity of nonstationary variables is sometimes regarded as a plausible alternative to a random walk. Usually, a joint significance test for the composite null H_0: $\rho_{j1}=\rho_{j2}=0$ is closely related to the problem of testing for a unit root against a deterministic trend and, therefore, can be used to discriminate between the difference-stationary (DS) and trend-stationary (TS) processes. If $\rho_{j1}=\rho_{j2}=0$, we have what is known as the DS process. In this case, it is preferable to work on first differences, as first-differencing will yield stationary series. On the other hand, if $\rho_{j1}<0$ and $\rho_{j2}\neq 0$, we have the TS process. In this case, it is desirable to work on levels, as detrending will be enough to make the series stationary.

In fact, if the order of magnitude of the dependent variable Y_{jt} is

to be the same both under the null hypothesis of a unit root and under the alternative of trend-stationarity, it is necessary that the test regression should include an unconstrained linear time trend with a coefficient that is zero $(\rho_{j2}=0)$ under the null $(\rho_{j1}=0)$ and nonzero $(\rho_{j2}\neq 0)$ under the alternative $(\rho_{j1}<0)$. For further details see Davidson and Mackinnon (1993, pp.709-10).

In any such situation, the standard regularity conditions do not hold, but one can still perform a t-test or an F-test provided the right critical values are used. Although the ADF tests are widely used, their discriminatory power can be quite limited, as a failure to reject the hypothesis of a unit root merely provides weak evidence that the variables are I(1). As West (1987) argues, the use of test regression (5.1a) (without a time trend) will be inconsistent if the true data process is stationary about a trend. By contrast, it has been shown by Perron (1989) that the use of test regression (5.1b) (with a time trend) may still tend to lack power if the true process is stationary about breaking levels or trends. Also, the danger with the formulation of model (5.1b) should be clear, as the forecasts made on the basis of an inappropriate deterministic trend may be very misleading. See, for a detailed discussion, Harvey (1992), Chapter 6.

5.2.1b *Testing for the Presence and Rank of Cointegration*

When we are satisfied that the time series are individually I(1), the next key step for those integrated variables is to detect the eventual presence of cointegration and to set the value of cointegrating rank, r, from available past observations. There are two main techniques involved in testing for cointegration: Engle and Granger's special ADF test (for r=1), and Johansen's ML procedure (for $0<r<K$). For comparative purposes, both approaches will be briefly examined for potentially cointegrated variables in a multivariate context.

(i) Engle and Granger's augmented Dickey-Fuller (ADF) test

The test is for whether there exists a cointegration relation, $\underline{\alpha}'\underline{Y}_t$

$=\alpha_1 Y_{1t}+\ldots+\alpha_K Y_{Kt}$, among the chosen set of I(1) variables. Consider a static OLS regression of one variable, say Y_1, arbitrarily normalized to have a unit coefficient upon the other variables

$$Y_{1t} = \eta_2 Y_{2t} + \ldots + \eta_K Y_{Kt} + \varepsilon_{1t}, \quad t = 1, \ldots, T, \qquad (5.2)$$

where $\eta_i=-\alpha_i/\alpha_1$, $\alpha_1\neq 0$. The variables are then cointegrated in the terminology of Engle and Granger (1987, pp.260-4) if, and only if, ε_{1t} is stationary or I(0). This single-equation regression will be called the cointegrating regression as the CV can be estimated efficiently without concern over the dynamics.

Specifically, Engle and Granger's suggested ADF statistic, sometimes called the augmented Engle-Granger test or the AEG test, can be performed in the following manner. First construct the cointegration residual $\hat{\varepsilon}_{1t}$ from (5.2) by least squares, and then test its stationarity by using unit root tests applied to $\hat{\varepsilon}_{1t}$

$$\nabla\hat{\varepsilon}_{1t} = \rho\hat{\varepsilon}_{1,t-1} + \sum_{i=1}^{s} \gamma_1 \nabla\hat{\varepsilon}_{1,t-i} + v_{1t},$$

where again s is selected to be sufficiently large to ensure that the residuals $\hat{v}_{1t}$ are serially uncorrelated. Here, the assumption of no cointegration, or spurious regression, is taken as the null hypothesis in the residual-based cointegration tests, mainly because it is consistent with the null of a unit root in the OLS residuals $(\hat{\varepsilon}_{1t})$ of the cointegration regression. The pseudo t-ratio for $\rho=0$ versus $\rho<0$ thus corresponds to the AEG test for the null of noncointegration versus the alternative of cointegration. In this sense, a one-sided statistic that is significantly negative in large samples would give rise to a rejection of the null of no cointegration.

If cointegration does exist, the residual $\hat{\varepsilon}_{1t}$ saved from the cointegration regression can then be fed as an additional regressor into a full dynamic error-correction system at each period. Otherwise, the model should be built on differences only. In cointegration tests, the required critical values for the AEG test are provided by MacKinnon (1990), using response surface estimation,

and differ from the critical values developed and used in tests for unit roots.

(2) Johansen's log-likelihood ratio (LR) test

In order to see how the empirical rank test proceeds for a model with cointegration rank r, $0<r<K$, consider a K-dimensional, cointegrated Gaussian VAR(p) process as in (4.5), perhaps with trend.[3] This model will be further reparameterised as in (4.9) to make it feasible that a number of potentially interesting economic hypotheses can be tested, i.e.:

$$\nabla \underline{Y}_t = \underline{C} + \sum_{j=1}^{p-1} \underline{A}_j^{\dagger} \nabla \underline{Y}_{t-j} - \underline{A}_p(1) \underline{Y}_{t-p} + \underline{u}_t, \quad t = 1, \ldots, T.$$

The cointegration constraint concerning the deficient rank of the long-run equilibrium matrix $\underline{A}_p(1) \left(= I_K - \sum_{i=1}^{p} \underline{A}_i\right)$, namely r, can be written as

$$H(r): \; rk\left[\underline{A}_p(1)\right] = r \quad \text{or} \quad \underline{A}_p(1) = \underline{\omega}\underline{\alpha}'. \tag{5.3}$$

Here $\underline{\alpha}$ and $\underline{\omega}$ are, as defined previously, Kxr cointegrating and loading matrices of full column rank r.

Under the hypothesis of cointegration, the Johansen-type log-likelihood ratio (LR) test statistic begins by running two auxiliary OLS regressions, with an intercept term included

$$\nabla \underline{Y}_t = \underline{C}_0 + \sum_{i=1}^{p-1} \underline{\Gamma}_{0i} \nabla \underline{Y}_{t-i} + \underline{R}_{0t},$$

and

$$\underline{Y}_{t-p} = \underline{C}_p + \sum_{i=1}^{p-1} \underline{\Gamma}_{pi} \nabla \underline{Y}_{t-i} + \underline{R}_{pt}.$$

Formulating the KxK product moment matrices of the estimation residuals

---

[3] For ease of discussion, we exclude a linear time trend from both the data generating process and the test formulation considered here. Inference on the presence of a deterministic trend can be conducted as pointed out in Johansen (1992) and Johansen & Juselius (1990, 1992).

$$\hat{\underline{S}}_{ij} = T^{-1}\sum_{t=1}^{T} \hat{\underline{R}}_{it}\hat{\underline{R}}_{jt}', \qquad i, j = 0, p,$$

the LR test of the null of there being at most r CVs, or equivalently K-r roots at, or close to, zero, is given by

$$\lambda_{LR}(r, K) = -T\sum_{i=r+1}^{K} \ln(1 - \hat{\lambda}_i), \tag{5.4}$$

where K is small relative to T, and $\hat{\lambda}_{r+1} \geq \ldots \geq \hat{\lambda}_K$ are the K-r smallest ordered eigenvalues of the determinantal equation

$$\left|\hat{\lambda}\hat{\underline{S}}_{pp} - \hat{\underline{S}}_{p0}\hat{\underline{S}}_{00}^{-1}\hat{\underline{S}}_{0p}\right| = 0. \tag{5.5}$$

Now suppose we want to test a specific, more restricted cointegration rank $r=r_0$ against a larger, less restricted rank of cointegration, say $r=r_1$ or $r=r_0+1$, in a sequential manner. In other words, we wish to test

$$H_0: r=r_0 \quad \text{against} \quad H_a: r_0 < r = r_1 \leq K,$$

and

$$H_0: r=r_0 \quad \text{against} \quad H_a: r = r_0 + 1.$$

From the cointegration LR test of (5.4), Johansen's (1988, 1991) trace statistic for testing $H(r_0)$ against $H(r_1)$ is then given by

$$\begin{aligned}\lambda_{tr}(r_0, r_1) &= T\left[-\sum_{i=r_0+1}^{K} \ln(1 - \hat{\lambda}_i) + \sum_{i=r_1+1}^{K} \ln(1 - \hat{\lambda}_i)\right] \\ &= -T\sum_{i=r_0+1}^{r_1} \ln(1 - \hat{\lambda}_i),\end{aligned} \tag{5.6}$$

and Johansen's (1989) maximum eigenvalue statistic for testing $H(r_0)$ against $H(r_0+1)$ is now given by

$$\lambda_{max}(r_0, r_0+1) = -T\ln(1 - \hat{\lambda}_{r_0+1}). \tag{5.7}$$

Both tests have nonstandard limit distributions which turn out to depend on just one parameter, K-r, and must be evaluated numerically. Certain asymptotic critical values for these two statistics have been tabulated by simulation in Johansen and

Juselius (1990) (K-r=1,...,5); these tables have been extended (K-r=1,...,10) by Osterwald-Lenum (1992). In addition, the λ_{max} test is directly comparable to Engle and Granger's residual-based ADF test in the case of r=1. This test is used in the empirical studies.

By employment of Johansen's λ_{max} test, the number of CVs can be determined using a series of log-likelihood ratio test statistics. In such a sequence, one way to proceed is to start with the null of no cointegration and work up. That is, should the null hypothesis that r=0 be rejected, we may then test the hypothesis that there is at most one CV ($r \leq 1$), and so forth until the null hypothesis fails to be rejected for the first time. The test results provide evidence in favour of cointegration only in the case where $0<r<K$.

In practical applications, there are three points to emphasize. Firstly, critical values of the asymptotic distributions of the LR statistics will vary when the system contains no constant, just a constant, and both a constant and a time trend. Secondly, the small sample critical values of the test statistics may differ slightly from the asymptotic critical values. In such cases, we should replace T by T-(Kp+1) in the above formula to adjust the test value in a small sample case [Reinsel & Ahn (1988)]. Finally, there seems to be a tendency to overestimate r when the true value is low and the costs of incorrectly specifying r are higher for overestimates [see Brandner & Kunst (1990)].

5.2.1c *Estimation of Cointegrating Vector(s)*

As a continuation, this section aims to derive the appropriate values of important CV(s) for a multivariate cointegrated system with independent Gaussian errors. For this purpose, the same two methods mentioned in the foregoing will be used and their relevant asymptotic properties will then be compared, given that the rank of the cointegration space, r, is known.

(1) When r=1

The estimates of the K free elements of the unique CV can be readily

obtained from equation (5.2) by OLS. This in turn motivates the 'two-stage estimator' proposed for the hypothesis of cointegration and its closely related ECM representation. That is, the estimator

$$\hat{\underline{\eta}} = \underline{Y}_{(1)}' \underline{y}_{(2)} \left(\underline{y}_{(2)}' \underline{y}_{(2)} \right)^{-1} \tag{5.8}$$

is used for $\underline{\eta}=(\eta_2,\ldots,\eta_K)$ in the first stage, and the remaining parameters are estimated conditionally on the estimator $\hat{\underline{\alpha}}'=(1,-\hat{\underline{\eta}})$ of $\underline{\alpha}'=(\alpha_1,\ldots,\alpha_K)$ in the second. Here

$$\underline{Y}_{(1)} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1T} \end{bmatrix} \quad \text{and} \quad \underline{y}_{(2)} = \begin{bmatrix} Y_{21} & \cdots & Y_{K1} \\ \vdots & & \vdots \\ Y_{2T} & \cdots & Y_{KT} \end{bmatrix}.$$

Stock (1987) has shown that when series are cointegrated, the OLS estimator $\hat{\underline{\eta}}$ is superconsistent.[4] More precisely, we have

$$\text{plim } T^{1-\delta}(\hat{\underline{\eta}} - \underline{\eta}) = 0 \qquad \text{for all } \delta>0.$$

The proposition implies that $\hat{\underline{\eta}}$ approaches the true value $\underline{\eta}$ at a faster rate proportional to T^{-1} than would be so with standard asymptotics, but $\hat{\underline{\eta}}$ is consistent with finite-sample bias.

If the disturbances ε_{1t} were actually observable, the unit root test statistics would have the same asymptotic distributions as the ones discussed previously. However, in almost all cases, ε_{1t} will not be observed and will therefore have to be estimated using OLS. The main reason that the asymptotic distributions of conventional unit root tests cannot be used in the present residual-based cointegration tests is that, if the null of noncointegration is true, the CV $\underline{\alpha}$ will not be identified and the disturbances ε_{1t} will have theoretically infinite variance. Heuristically, OLS will seek the vector to minimize the residual variance and is very likely to make the estimated residuals stationary. This implies that the ordinary unit root test statistics will reject the null too often if $\underline{\alpha}$ must be estimated. To avoid this possibility, the related critical values have to be raised accordingly.

[4] Even so, standard inference procedures still do not apply.

Although conventional regression estimates may be of considerable importance in applied research, many objections have been raised as to the adequacy of this method in the light of substantial small sample biases. The validity of its underlying assumptions has also been questioned. Not only is the number of CVs assumed to be unique before estimation, but the element of CV on arbitrary normalization is also known to be nonzero ahead of time. These are strong assumptions which may be somewhat unrealistic, especially when there are more than two I(1) variables under consideration. The regression estimator, however, provides no framework for addressing these issues. Nor does the estimator have well defined limiting distributions and, as a result, testing for cointegration is not a straightforward procedure [cf. Hall (1989), p.213].

It, therefore, appears that the maximum likelihood (ML) estimation procedure put forward by Johansen (1988, 1991) and Johansen and Juselius (1990, 1992, 1994) may be more satisfactory. It relaxes the assumption that the CV is unique, provides a unified framework for the estimation and testing of one or more cointegrating relations, and also takes into account the error structure of the underlying process. Therefore it may be expected to behave better than the traditional regression estimates.

(2) When $0<r<K$

To implement the Johansen ML procedure, we denote the K orthonormalized eigenvectors of $\hat{\underline{S}}_{p0}\hat{\underline{S}}_{00}^{-1}\hat{\underline{S}}_{0p}$ with respect to $\hat{\underline{S}}_{pp}$ as a $K\times K$ matrix $\hat{\underline{V}}=(\hat{\underline{V}}_1,\ldots,\hat{\underline{V}}_K)$ corresponding to the K nonincreasingly ordered eigenvalues $\hat{\lambda}_1\geq\ldots\geq\hat{\lambda}_K$, and normalize $\hat{\underline{V}}$ by $\hat{\underline{V}}'\hat{\underline{S}}_{pp}\hat{\underline{V}}=I_K$. All K eigenvalues and eigenvectors are permitted to be calculated at once. Then the ML estimators of the space spanned by $\underline{\alpha}$, the cointegration space $sp(\underline{\alpha})$, and the space spanned by $\underline{\omega}$, the adjustment space $sp(\underline{\omega})$, are given, respectively, by

$$\hat{\underline{\alpha}} = (\hat{\underline{V}}_1, \ldots, \hat{\underline{V}}_r), \tag{5.9a}$$

and

$$\hat{\underline{\omega}} = -\hat{\underline{S}}_{0p}\hat{\underline{\alpha}}. \tag{5.9b}$$

The r mutually independent columns of $\hat{\underline{\alpha}}$, i.e. the first r columns of $\hat{\underline{V}}$ that correspond to the r largest eigenvalues, constitute a basis of the r-dimensional cointegration space, as all possible choices of the optimal $\underline{\alpha}$ can be derived from such a basis. Moreover, every single row of the impact matrix $\underline{A}_p(1)$ can be thought of as a weighted average of the r linearly independent columns of $\hat{\underline{\alpha}}$, while every single column of $\underline{A}_p(1)$ can be regarded as a weighted average of the r columns of $\hat{\underline{\omega}}$. Once a specific choice is made for $\underline{\alpha}$, such as $\underline{\alpha}=\hat{\underline{\alpha}}\underline{P}$, for $\underline{P}$ a nonsingular r×r matrix, $\underline{\omega}$ will subsequently be uniquely determined as $\underline{\omega}=\hat{\underline{\omega}}\underline{P}'^{-1}$. This is important because the apparent problem that the regression model of (4.9) depends nonlinearly on the parameters can be circumvented [cf. Johansen (1988) and Davidson & Mackinnon (1993), Chapter 20].

It should perhaps be emphasised that what can be determined by the model is only a basis of the space spanned by $\hat{\underline{\alpha}}$, $sp(\underline{\alpha})$, and of the space spanned by $\underline{\omega}$, $sp(\underline{\omega})$, which can be estimated superconsistently. Yet the K×r matrices $\underline{\alpha}$ and $\underline{\omega}$ themselves are not identified by virtue of being not unique and cannot be estimated consistently at this stage. However, conditional on these bases, the long-run matrix $\underline{A}_p(1)$ as well as the other parameters in the system (4.9) can be estimated consistently without further identifying constraints. The identification of the different cointegrating relations and the associated adjustment coefficients, on the other hand, has to be done *a posteriori* by imposing linear restrictions on either the cointegration space or the adjustment space [see Urbain (1995)]. In this procedure, the normality of the innovations in (4.9) is required to ensure asymptotically efficient ML estimators of the CVs and the adjustment coefficients. The robustness of the procedure to departures from normality is as yet unknown [see Johansen & Juselius (1990), p.176].

When the cointegrating relations have been determined in the sense that several subsets of the I(1) variables tend to bunch together in the long run, the next logical step is to seek an appropriate prior structure for the coefficients of the multivariate cointegrated system under study.

5.2.2 Determination of Prior Restrictions

So far, we have described the determination of the plausible long-run restrictions on the matrix $\underline{A}_p(1)$, but the other parameters $\underline{A}_j^{\dagger}$, j=1,...,p-1, have no constraints imposed and would be treated as equally likely. This position may lead to a distortion of the modeller's beliefs concerning the best forecasting model, even though the VAR-ECM is estimated as if not subject to any restrictions other than (5.3). Given a finite number of observations, the bigger the model is, the weaker will be the data evidence on the individual coefficients and, hence, the more important the stochastic prior regarding the distributions of the coefficients. The importance of the prior in potentially improving the precision of estimates may lead to a consideration of whether an explicitly Bayesian procedure might produce better forecasts than those generated by other competing approaches. The BVAR forecasting technique has thus been developed to test this possibility.

Here, the prior restrictions for the parameter estimates will be specified in the Bayesian prior mean-variance framework. Under the Minnesota or Litterman prior, the prior distributions for all but coefficients of the deterministic components are assumed to be independent normal, so that they can be uniquely determined by just two prior parameters — namely the prior means and the prior variances. The prior covariances are always set at zero.[5] Generally, the standard Minnesota prior is implemented by specifying across all equations the same linear form of prior distribution and placing on the high order lags a tight zero prior with small standard deviations. However, the appearance of the non-zero prior standard deviations implies that we are not sure about such a simple model, compatible with the uncertainty that the modeller is prepared to

[5] In principle, it is possible and also important that such covariances between parameters, across different equations as well as in the same equation, are allowed for, as they may be of value in improving forecasts. However, this goes beyond the scope of our present work. For a useful survey of this literature, see the recent articles by Holden (1995), Chapter 1, Kadiyala & Karlsson (1993, 1997), Karlsson & Lütkepohl (1993) and Westlund (1993).

allow for. Also, uncertainty of the prior standard deviation can be formally expressed by assigning to that parameter a further prior distribution, which is often shaped as a rectangular distribution confined to a certain range of values and governed by a scale factor called a prior hyperparameter. The priors placed on the deterministic terms are supposed to be 'flat', or noninformative, reflecting our complete ignorance about these coefficients.

The data sample will then be examined, using a standard Bayesian statistical procedure, to revise each of the modeller's best guesses (the prior mean) in terms of the modeller's initial confidence (the prior standard deviation). The larger the prior standard deviation, the more weight will be given to the accidental patterns in the data and vice-versa. Moreover, along with the design of prior structure, the potential influence of estimated cointegration relations on forecast accuracy will also be investigated.

5.2.2a *A Simple BVAR with no EC Model*

Consider first a K-dimensional Gaussian VAR(p) model of the form of (4.5)

$$\underline{Y}_t = \underline{C} + \sum_{i=1}^{p} \underline{A}_i \underline{Y}_{t-i} + \underline{u}_t; \quad t = 1, \ldots, T, \tag{5.10a}$$

or

$$Y_{jt} = C_j + \sum_{n=1}^{K} \sum_{i=1}^{p} A_{jn,i} Y_{n,t-i} + u_{jt}; \tag{5.10b}$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \quad j = 1, \ldots, K; \quad t = 1, \ldots, T.$$

The prior restrictions concerning the possible values of the parameters $\underline{A}_i$, i=1,...,p, can be set up in the fashion of Litterman (1986a, 1986c), as follows.

(1) The prior mean: a multivariate random walk process

In the standard Minnesota prior, the means of the prior distributions for all coefficients are usually set to zero, except the first own lag in each equation, which has a prior mean of unity by default. This prior mainly reflects the assumption that most of

the variation in each of the variables is accounted for by own recent lags and loosely centres about a vector random walk process ($p=1$, $\underline{A}_1=I_K$ and $\underline{C}=0$)

$$\underline{Y}_t = \underline{Y}_{t-1} + \underline{u}_t. \tag{5.11a}$$

Alternatively, we may consider a vector random walk plus drift process ($p=1$, $\underline{A}_1=I_K$ and $\underline{C}\neq 0$)

$$\underline{Y}_t = \underline{C} + \underline{Y}_{t-1} + \underline{u}_t. \tag{5.11b}$$

These seem to be reasonable approximations for typically behaved macroeconomic variables that may appear to exhibit relatively smooth random walk components. If this is so, the best forecast would be produced essentially by a naive 'no-change' prediction. That is, in the case in which the drift parameter vector $\underline{C}=0$, future values of the variables would differ from their current values only because of completely unpredictable random events. The inclusion of drift $(\underline{C}\neq 0)$ unrestrictedly in all of the equations is simply to represent an explicit steady-state rate or growth per period in the variables being modelled. In practice, even for some variables whose changes are thought to be partially predictable, this naive forecasting approach can be surprisingly difficult to improve upon.

Given the prior means, all that is required is to specify their standard deviations so as to complete the determination of the prior distributions.

(2) The prior standard deviation

In order to construct standard deviations of the priors for the model's successive lag coefficients, the estimation of a set of auxiliary unrestricted univariate autoregressions (ARs) is first conducted on the basis of T observations

$$Y_{jt} = c_j + \sum_{i=1}^{p} a_{ji} Y_{j,t-i} + e_{jt}, \tag{5.12}$$

$$e_{jt} \sim \text{NID}(0, \sigma_j^2), \quad t = 1, \ldots, T, \quad j = 1, \ldots, K.$$

The estimated standard errors $\hat{\sigma}_j$, $j=1,\ldots,K$, of the residuals are

saved for all variables. This is because normally the prior standard deviations around coefficients on lags of the cross variables in each equation of the BVAR model are not scale invariant. To adjust for the scale variability of the actual data, the BVAR prior must be specified with reference to the relative sizes of unexpected movements in different series. Such movements are often reflected in the standard errors of the residuals computed from linear regressions of the series on several of their own past values.

The standard deviation of the distribution for each coefficient $A_{jk,l}$, the jk-th element of $\underline{A}_l$, in the system through use of a general prior — which is largely the same as the symmetric — for all j, k, l: denoted s(j,k,l), can then be written as

$$s(j,\ k,\ l) = \nu f(j,\ k) g(l) \hat{\sigma}_j / \hat{\sigma}_k,$$
$$j,\ k = 1,\ \ldots,\ K, \quad l = 1,\ \ldots,\ p.$$

Here the ratio $\hat{\sigma}_j / \hat{\sigma}_k$ is involved as a rescaling factor, which converts the cross lag standard deviations to units comparable to those of the own lag standard deviations in equation j. The diagonal matrix $\underline{\nu}$ is the overall tightness of the prior, which directly controls the standard deviation on the first own lag of the normalized variable of each equation. The function g(l) is the lag pattern of (harmonic) decay, which increasingly restricts the influence of past values on recent ones. The function f(j,k) is the general prior type, which simultaneously governs the weights of all off-diagonal elements relative to the weight of the diagonal element in each row of $\underline{A}_l$.

To be more specific, we have

$$S(j,\ k,\ l) = \begin{cases} \nu_j / l & \text{if } j = k \\ \nu_j w_{jk} \hat{\sigma}_j / l \hat{\sigma}_k & \text{if } j \neq k, \end{cases} \tag{5.13}$$

or in matrix form,

$$S(j,\ k,\ l) = \begin{bmatrix} \nu_1/l & \nu_1 w_{12} \hat{\sigma}_1 / l \hat{\sigma}_2 & \cdots & \nu_1 w_{1K} \hat{\sigma}_1 / l \hat{\sigma}_K \\ \vdots & \ddots & & \vdots \\ \nu_j w_{j1} \hat{\sigma}_j / l \hat{\sigma}_1 & & \nu_j / l & \nu_j w_{jK} \hat{\sigma}_j / l \hat{\sigma}_K \\ \vdots & & \ddots & \vdots \\ \nu_K w_{K1} \hat{\sigma}_K / l \hat{\sigma}_1 & \cdots & \nu_K w_{K,K-1} \hat{\sigma}_K / l \hat{\sigma}_{K-1} & \nu_K / l \end{bmatrix}_{(K \times K)}.$$

Here, the scale factor or element ν_j or $\nu_j w_{jk}$, $j,k\in[1,K]$, in the rows is the particular hyperparameter assigned to each of the K variables in each of the K equations of the system individually. According to the modeller's prior beliefs, the parameters that determine the prior standard deviations are frequently confined to a certain narrow range, say $0<\nu_j<3$ and $0<w_{jk}<1$, and are regarded as equally likely within the range, but highly unlikely outside it. Again, during the modelling process, it is important to remember that the Bayesian procedure of Litterman does not take account of dependencies between VAR coefficients and is thus of a univariate nature. Hence, in this analysis, whether the overall tightness (ν_j) and the fraction of tightness (w_{jk}) are set to too 'large' or 'small' values over the finite range will exert a direct impact on the type of the model that one should construct. In particular, three special cases need to be considered.

Case (a): $\nu_j=3$ and $w_{jk}=1$ — that is, both ν_j and w_{jk} are set to be large values. In this case, the 'Bayesian' part of the BVAR will virtually be cut out. The model will approach an UVAR model and the data will dominate determination of the coefficients almost completely. Hence, if the variables are believed to be closely related, a large value for w_{jk} is acceptable.

Case (b): $\nu_j=3$ and $w_{jk}=0$ — that is, ν_j is set to be a large value while w_{jk} is zero. As such, the 'Vector' part of the BVAR will essentially be chopped off. The system will reduce to a set of K unrestricted univariate autoregressions and the modeller's chance of discovering any important but unexpected historical relationships between variables will be sharply limited. Thus, if the variables are expected to be loosely related, a small value for w_{jk} is reasonable.

Case (c): $\nu_j=0$ — that is, ν_j is set to be zero whatever w_{jk} is. In this case, the system will eventually shrink to the standard Minnesota prior means — K independent random walks and the prior will dictate the coefficients completely. Hence, if the variables are thought to be roughly unrelated random walks, a small value for ν_j is appropriate.

These three cases represent extremes and the explicitly BVAR forecasting model will most likely take intermediate values within the specified ranges. Actually, in any intermediate position, holding ν_j and increasing w_{jk} will mean that the standard deviation of the prior will tend to be about the same for the coefficients of each row of $\underline{A}_1$ except for a scaling factor $(\hat{\sigma}_j/\hat{\sigma}_k)$ that takes care of different magnitudes of different variables. Holding ν_j and lowering w_{jk} will force all the off-diagonal elements, at all lags, towards zero. On the other hand, holding w_{jk} and increasing ν_j will tend to eliminate the effect of the prior on all coefficients, whereas holding w_{jk} and tightening ν_j will push the diagonal elements of $\underline{A}_1$ towards 1 and all the off-diagonal elements towards zero. In all cases, the intercept terms are left unrestricted and assigned an effectively infinite prior standard deviation.

The estimated univariate residual variance $\hat{\sigma}_j^2$ described above can be replaced approximately by the j-th diagonal element $\hat{\sigma}_{jj}^2 (\leq \hat{\sigma}_j^2)$ of the K×K estimated residual covariance matrix $\tilde{\Sigma}_u$ of the multivariate system (5.10a) if, and only if, the values of ν_j and w_{jk} are chosen to be 3 and 0, respectively.[6]

In searching for the most appropriate setting of the hyperparameter values, the objective facing the forecaster is to select that set of the hyperparameters which allows a forecast to be made optimally. For this, one usual but informal method for picking among plausible values is to see how a model based on alternative hyperparameters would have within-sample forecasts, provided that the same regression process will continue to hold over the future period. Various moderate weights are tried on those aggregates which are

---

[6] In system (5.10a), the resulting ML estimator of Σ_u can be expressed as:

$$\tilde{\Sigma}_u = \underset{(K\times T)}{\underline{Y}}\left(I_T - \underline{X}'(\underline{X}\underline{X}')^{-1}\underline{X}\right)\underline{Y}'/T,$$

where $\underline{Y}=[\underline{Y}_1,\ldots,\underline{Y}_T]$, $\underline{X}=[\underline{X}_0,\ldots,\underline{X}_{T-1}]$, and $\underline{X}_t=\begin{bmatrix}1\\ \underline{Y}_t\\ \vdots\\ \underline{Y}_{t-p+1}\end{bmatrix}$.

$((1+Kp)\times T)$ $\quad$ $((1+Kp)\times 1)$

believed to be important for the prediction of the own variable in question, while the weights on others which are expected to be less important are held low and fixed. The best setting — and thus the best final forecasting model — would be the one that seems to lead to the smallest average forecasting errors of the historical data.

The j-th equation of the system with the priors presented above can thus be written as

$$
\begin{aligned}
Y_{jt} = {} & \underset{(\infty)}{0} + \underset{(\nu_j w_{j1} \hat{\sigma}_j / \hat{\sigma}_1)}{0}\, Y_{1,t-1} + \ldots + \underset{(\nu_j w_{j1} \hat{\sigma}_j / p\hat{\sigma}_1)}{0}\, Y_{1,t-p} + \ldots \\
& + \underset{(\nu_j)}{1}\, Y_{j,t-1} + \ldots + \underset{(\nu_j / p)}{0}\, Y_{j,t-p} + \ldots \\
& + \underset{(\nu_j w_{jK} \hat{\sigma}_j / \hat{\sigma}_K)}{0}\, Y_{K,t-1} + \ldots + \underset{(\nu_j w_{jK} \hat{\sigma}_j / p\hat{\sigma}_K)}{0}\, Y_{K,t-p} + u_{jt};
\end{aligned}
$$

$$
u_{jt} \sim \mathrm{NID}(0,\ \sigma_{jj}^2); \quad 0 < \nu_j \leq 3; \quad 0 < w_{jk} \leq 1;
$$

$$
j,\ k = 1,\ \ldots,\ K; \quad t = 1,\ \ldots,\ T,
$$

where all coefficients are set at their prior means and their prior standard deviations are given in parentheses. In addition, as in the stationary case, the prior parameter covariance matrix of this typical equation may be reformulated as a $(1+Kp)\times(1+Kp)$ diagonal matrix; that is,

$$
\begin{aligned}
\underline{V}_j = \mathrm{diag}[\ & \infty,\ (\nu_j w_{j1} \hat{\sigma}_j / \hat{\sigma}_1)^2,\ \ldots,\ (\nu_j w_{j1} \hat{\sigma}_j / p\hat{\sigma}_1)^2,\ \ldots, \\
& (\nu_j)^2,\ \ldots,\ (\nu_j / p)^2,\ \ldots, \\
& (\nu_j w_{jK} \hat{\sigma}_j / \hat{\sigma}_K)^2,\ \ldots,\ (\nu_j w_{jK} \hat{\sigma}_j / p\hat{\sigma}_K)^2\ ].
\end{aligned}
$$

Using this reformulation, a more general form of the prior for the system will be provided, and various problems associated with it will be further discussed.

(3) A general expression of the prior

A general expression of the prior about the model's mean parameter values and their standard deviations can be given by graphical and algebraic methods respectively, as illustrated below.

(a) A graphical display

The combined effect of fixed prior means and varying prior standard deviations for the lag coefficients of the typical equation j is demonstrated in the accompanying figure 5.1.

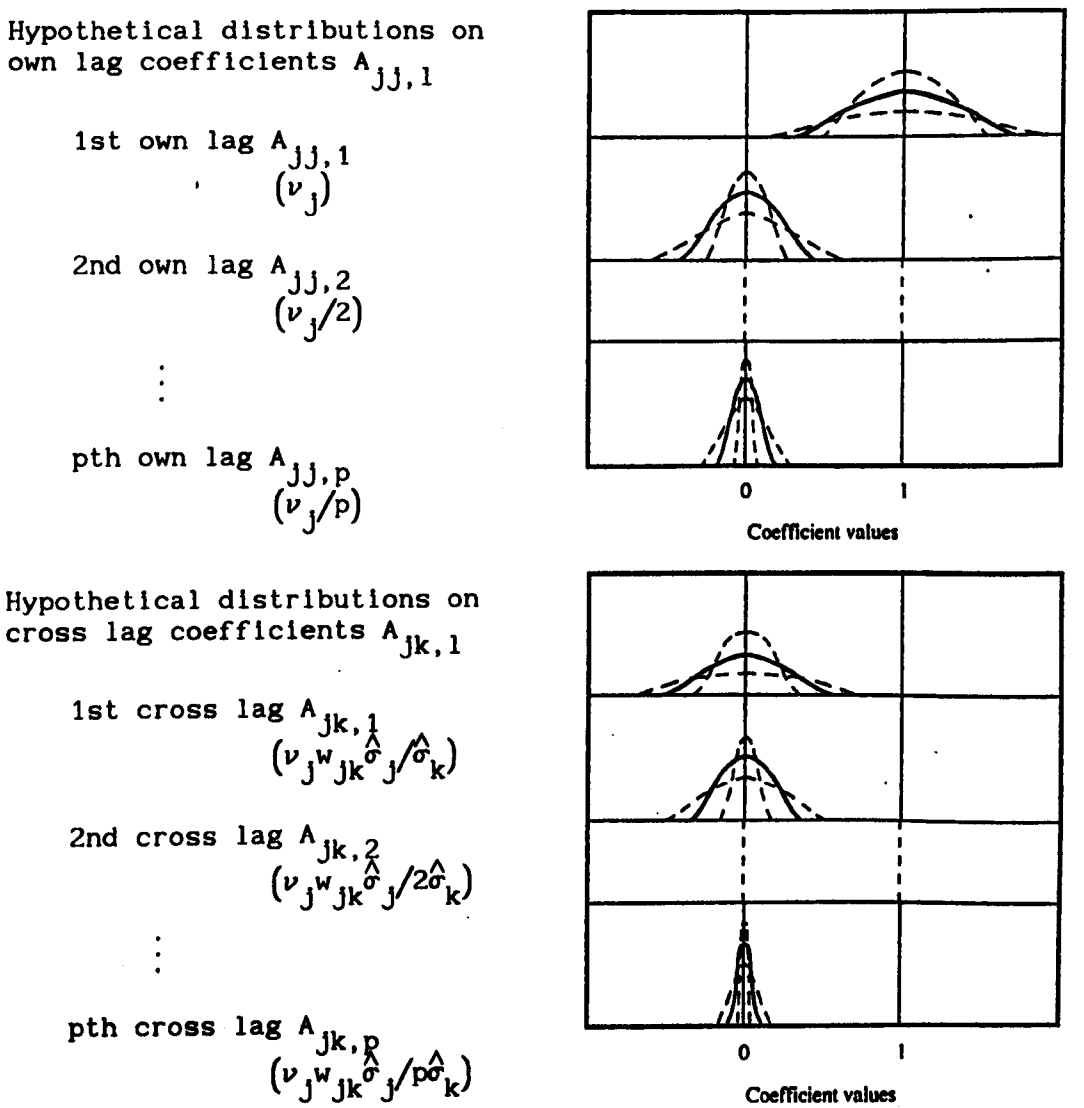

Fig. 5.1 The loose random walk prior on the lag coefficients of the typical equation j.

Fig. 5.1 shows two main kinds of normal curves attached to the coefficients of the own and cross lags in a typical equation.

With regard to the prior means, no matter how the standard deviations change within the limits, the curves for all lag coefficients other than the first own lag have their highest values

(the highest prior densities over the regions for which they are defined) at zero. The curve for the first lag of the own variable has its highest value at one.

With respect to the prior standard deviations, though they are each allowed to vary in value, the curves for recent values of variables are more likely to be broad and low than less recent ones. Also, regardless of the scale variability of different variables, the curves for the own lags are more likely to be broad and low than the corresponding cross ones. Hence, the curve for the first own lag is relatively the broadest and lowest, which means that a wide range of possible values for the coefficient has density not much lower than the best guess, and that even values fairly far from the best guess are not considered to be extremely improbable.

As lag length l increases, the curves for further lags of variables become progressively more peaked and concentrated around the best guess, restricting values to some narrow ranges. This implies that with increasing lag length the modeller becomes more confident that a zero coefficient will be consistent with a model that is likely to forecast well.

(b) An algebraic description

The chosen stochastic prior information on the parameter vector $\underline{\beta}_j$ can be represented by

$$\underline{r}_j = \underline{R}_j\underline{\beta}_j + \underline{v}_j \quad \text{with } \underline{v}_j \sim \text{NID}(0,\ I_{1+Kp}); \qquad (5.14)$$
$$j = 1, \ldots, K,$$

where

$$\underline{r}_j \triangleq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \underline{R}_j \triangleq \begin{bmatrix} 0 & & & & & & 0 \\ & \hat{\sigma}_1/\nu_j w_{j1}\hat{\sigma}_j & & & & & \\ & & \ddots & & & & \\ & & & 1/\nu_j & & & \\ & & & & \ddots & & \\ & & & & & p/\nu_j & \\ & & & & & & \ddots \\ 0 & & & & & & p\hat{\sigma}_K/\nu_j w_{jK}\hat{\sigma}_j \end{bmatrix}, \quad \underline{\beta}_j \triangleq \begin{bmatrix} C_j \\ A_{j1,1} \\ \vdots \\ A_{jj,1} \\ \vdots \\ A_{jj,p} \\ \vdots \\ A_{jK,p} \end{bmatrix}, \quad \underline{v}_j \triangleq \begin{bmatrix} v_{j0} \\ v_{j1,1} \\ \vdots \\ v_{jj,1} \\ \vdots \\ v_{jj,p} \\ \vdots \\ v_{jK,p} \end{bmatrix}.$$

$((1+Kp)\times 1)$ $\quad ((1+Kp)\times(1+Kp))$ $\quad ((1+Kp)\times 1)$ $\quad ((1+Kp)\times 1)$

In this representation, there is a single '1' in the known column vector $\underline{r}_j$ corresponding to the first lag of the dependent variable and a '0' on the diagonal of the known diagonal matrix $\underline{R}_j$ corresponding to the deterministic component. Therefore, the singularity of $\underline{R}_j\underline{R}_j'(=\underline{V}_j^{-1})$ (the inverse of the prior covariance matrix) simply reflects the improper flat prior imposed.[7]

The use of the BVAR model in levels takes account of nonstationarity in time series by imposing a unit prior on the first own lag. Symmetric distributions around this prior mean will inevitably allow for systems with undesirable explosive roots however. Although we may recognize that the standard Minnesota prior is not the best, we can also doubt that the gain that could be achieved by abandoning the Gaussian form for our prior would be worth the price [see, for details, Doan *et al.* (1984), p.7]. Also, when data are finite and nonstationary, it may well be preferable to impose the cointegration restrictions together with the prior restrictions. However, in this basic BVAR model, the information regarding cointegrating linear combinations may be incorporated only implicitly. Therefore, when approaching time series, a more sophisticated BVAR model that explicitly brings in both prior and cointegration restrictions is needed. A BVAR-EC model enables us to do just this.

5.2.2b *A Combined BVAR with EC Model*

Consider now a cointegrated Gaussian, K-dimensional VAR(p) model written in the error-correction version like (4.9), i.e.,

$$\nabla\underline{Y}_t = \underline{C} + \sum_{i=1}^{p-1} \underline{A}_i \nabla\underline{Y}_{t-i} - \underline{\omega}\underline{\varepsilon}_{t-p} + \underline{u}_t; \text{ with} \qquad (5.15a)$$

$$\underline{\varepsilon}_{t-p} = \underline{\alpha}'\underline{Y}_{t-p}; \quad t = 1, \ldots, T,$$

or

$$\nabla Y_{jt} = C_j + \sum_{n=1}^{K}\sum_{i=1}^{p-1} A_{jn,i} \nabla Y_{n,t-i} - \sum_{s=1}^{r} \omega_{js}\varepsilon_{s,t-p} + u_{jt}; \text{ with} \qquad (5.15b)$$

---

[7] Here, the use of $\underline{R}_j\underline{R}_j'$ may help avoid inconvenient algebraic operations with the element, infinity, in $\underline{V}_j$.

$$\varepsilon_{s,t-p} = \underline{\alpha}_s' \underline{Y}_{t-p} = \sum_{n=1}^{K} \alpha_{sn} Y_{n,t-p};$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \; 1 \le r \le K-1; \; j = 1, \ldots, K; \; t = 1, \ldots, T.$$

Both the prior and cointegration restrictions are explicitly imposed onto the parameters $\underline{A}_j^{\dagger}$, j=1,...,p-1, and $\underline{A}_p(1)$ for this system, which can be used to describe, explain and forecast the statistical variations of the data vector.

(1) The prior mean: a multivariate white noise process

From the derivation of $\underline{A}_j^{\dagger}(=-I_K+\sum_{i=1}^{j}\underline{A}_i)$, j=1,...,p-1, it can be seen that in accordance with the random walk hypothesis for the original VAR(p) coefficients $(\underline{A}_i)$, a prior mean of one for the first own lag will become a prior mean of zero after these linear transforms and all other coefficients have priors centred on zero. Moreover, the prior pdf for all lag coefficients $(\underline{A}_j^{\dagger})$ is still in the multivariate normal form (with mean zero and standard deviations nonzero). This is because $\underline{A}_j^{\dagger}$ are some linear combination of $\underline{A}_i$, all of which are assumed independently normally distributed. This prior is consistent not only with the random walk hypothesis for the nonstationary case but also with the white noise proposition for the stationary one. However, the error-correction term is negatively fed back into the system as an exogenous vector at full weight along with the constant term and the residual covariance matrix Σ_u is common to both competing models. In this situation, if the prior means were the true parameter values, the process would turn out to be a weighted sum of error-correction feedbacks and a multivariate white noise (with or without drift) process; that is,

$$\nabla \underline{Y}_t = - \underline{\omega\varepsilon}_{t-p} + \underline{u}_t, \qquad (5.16a)$$

or

$$\nabla \underline{Y}_t = \underline{C} - \underline{\omega\varepsilon}_{t-p} + \underline{u}_t. \qquad (5.16b)$$

This position seems to be reasonable for many macroeconomic variables, as most of them are believed to be I(1), and thus stationary upon first differencing and cointegrating linear

combinations. If this be the case, the real driving force that sets the system in motion must be the long-run equilibrium among the variables being modelled. The future movements of the data should therefore depend mainly on the useful long-run relationships, while the short-run fluctuations are regarded as fairly uninteresting.

Given the above argument, all that remains is to specify the prior parameter standard deviations that complete the determination of the prior distributions.

(2) The prior standard deviation

In order to facilitate the determination of the standard deviations for the model's lag coefficients, we start, as before, with the estimation of a set of auxiliary AR-EC equations

$$\nabla Y_{jt} + \hat{\underline{\omega}}_j \hat{\underline{\varepsilon}}_{t-p} = c_j^{\dagger} + \sum_{j=1}^{p-1} a_{ji}^{\dagger} \nabla Y_{j,t-i} + e_{jt}^{\dagger}; \qquad (5.17)$$

$$e_{jt}^{\dagger} \sim NID\left(0,\ \sigma_j^{\dagger 2}\right); \quad t = 1, \ldots, T; \quad j = 1, \ldots, K,$$

and store the estimated standard errors $\hat{\sigma}_j^{\dagger}$, $j=1,\ldots,K$, of the residuals for all differenced series. Here, an estimated error-correction term $\hat{\underline{\omega}}_j \hat{\underline{\varepsilon}}_{t-p}$ is included in the LHS with $\hat{\underline{\omega}}_j$, the j-th row vector of $\hat{\underline{\omega}}$, and $\hat{\underline{\varepsilon}}_{t-p}\left(=\hat{\underline{\alpha}}'\underline{Y}_{t-p}\right)$, the whole cointegration errors, since a pure AR representation on first differences is inappropriate in the presence of cointegration. However, there seems to be no reason why the residual standard errors measured above should be different from those obtained in (5.12).[8] Therefore, once we get the estimated residual standard errors $\left(\hat{\sigma}_j\right)$ in levels, they can be used in either case. The relative scales of these residual standard errors will then be considered for each lag of each cross

[8] As the univariate AR on (5.12) (or AR-EC on (5.17)) is equivalent to the multivariate VAR on (5.10a) (or VAR-EC on (5.15a)) with shrunken diagonal parameter matrices $\underline{A}_1$ (or $\underline{A}_j^{\dagger}$) and Σ_u is common to both models, the estimated residual standard errors obtained either from (5.12) in levels or from (5.17) in first differences should be treated as exactly the same; i.e., $\hat{\sigma}_j^{\dagger} \equiv \hat{\sigma}_j$ or $e_{jt}^{\dagger} \equiv e_{jt} \sim NID\left(0, \sigma_j^2\right)$.

variable in every equation, to make the standard deviations of different variables comparable.

The prior standard deviations for the elements $A^{\dagger}_{jk,l}$ of $\underline{A}^{\dagger}_{l}$, l=1,...,p-1, can then be formulated in a general prior type. Specifically, we have

$$S(j, k, l) = \begin{cases} \nu_j/l & \text{if } j = k \\ \nu_j w_{jk} \hat{\sigma}_j / l \hat{\sigma}_k & \text{if } j \neq k, \end{cases}$$

or, in matrix form,

$$S(j, k, l) = \begin{bmatrix} \nu_1/l & \nu_1 w_{12} \hat{\sigma}_1/l\hat{\sigma}_2 & \cdots & \nu_1 w_{1K} \hat{\sigma}_1/l\hat{\sigma}_K \\ \vdots & \ddots & & \vdots \\ \nu_j w_{j1} \hat{\sigma}_j/l\hat{\sigma}_1 & & \nu_j/l & \nu_j w_{jK} \hat{\sigma}_j/l\hat{\sigma}_K \\ \vdots & & \ddots & \vdots \\ \nu_K w_{K1} \hat{\sigma}_K/l\hat{\sigma}_1 & \cdots & \nu_K w_{K,K-1} \hat{\sigma}_K/l\hat{\sigma}_{K-1} & \nu_K/l \end{bmatrix}_{(K \times K)}.$$

Here again the overall tightness of the prior ν_j and its various fractions w_{jk} are allowed to take on any value in some particular intervals, say $0<\nu_j<3$ and $0<w_{jk}<1$, with uniform densities. In a search for values for the best setting of the hyperparameters, a similar method could be applied using an appropriate selection criterion (e.g. Theil-U statistic). The setting with the lowest average Theil-U values, and thus the best simulated out-of-sample forecasts within the sample period, would then be selected as the 'best'. To illustrate a composite effect of the choice of these prior standard deviation parameters and the hypothesis of $A_p(1)$ having a reduced rank r, $0<r<K$, the contents of Table 5.1 may be considered.

Table 5.1: The joint effect of the selection of the parameters ν_j, w_{jk} and r on the cointegrated Gaussian VAR(p) process

ν_j \ w_{jk} \ r	r=K	r=0	0<r<K
ν_j=3, w_{jk}=1	UVAR on levels	UVAR on first differences	VAR-ECM
ν_j=3, w_{jk}=0	K unrestricted univariate AR processes on levels	K unrestricted univariate AR processes on first differences	Unrestricted AR-EC processes
ν_j=3, 0<w_{jk}<1*	BVAR on levels with data dependent diagonal elements of $\underline{A}_1$	BVAR on first differences with data dependent diagonal elements of $\underline{A}_1^{\dagger}$	BVAR-ECM with data dependent diagonal elements of $\underline{A}_1^{\dagger}$
ν_j=0	K independent random walk processes	K independent white noise processes	EC feedbacks plus white noise processes
0<ν_j<3, w_{jk}=1*	BVAR on levels with (regardless of the scale variability) about the same prior standard deviation on each row elements of $\underline{A}_1$	BVAR on first differences with (irrespective of the scale variability) about the same prior standard deviation on each row elements of $\underline{A}_1^{\dagger}$	BVAR-ECM with (ignoring the scale variability) about the same prior standard deviation on each row elements of $\underline{A}_1^{\dagger}$
0<ν_j<3, w_{jk}=0	K restricted univariate AR processes on levels	K restricted univariate AR processes on first differences	Restricted AR-EC processes
0<ν_j<3, 0<w_{jk}<1*	BVAR on levels	BVAR on first differences	BVAR-ECM

* indicates various possible BVAR and BVAR-EC models.

In the BVAR-EC system studied here, the estimated error-correction term is often rather insensitive to the precise form of the prior structure adopted. The residual variance $\hat{\sigma}_j^{\dagger 2}$ (or $\hat{\sigma}_j^2$) can be substituted by the j-th diagonal element of $\tilde{\Sigma}_u$ of the system (5.15a) if, and only if, the values of ν_j and w_{jk} are taken to be 3 and 0, respectively.[9]

A typical equation of the system with both the prior and cointegration restrictions specified can thus be expressed as

$$\begin{aligned}
\nabla Y_{jt} = {} & 0 + \underset{(\infty)}{0} \nabla Y_{1,t-1} + \ldots + \underset{(\nu_j w_{j1} \hat{\sigma}_j / \hat{\sigma}_1)}{0} \nabla Y_{1,t-p+1} + \ldots \\
& + \underset{(\nu_j)}{0} \nabla Y_{j,t-1} + \ldots + \underset{(\nu_j/(p-1))}{0} \nabla Y_{j,t-p+1} + \ldots \\
& + \underset{(\nu_j w_{jK} \hat{\sigma}_j / \hat{\sigma}_K)}{0} \nabla Y_{K,t-1} + \ldots + \underset{(\nu_j w_{jK} \hat{\sigma}_j / (p-1) \hat{\sigma}_K)}{0} \nabla Y_{K,t-p+1} \\
& - \omega_{j1} \varepsilon_{1,t-p} - \ldots - \omega_{jr} \varepsilon_{r,t-p} + u_{jt};
\end{aligned}$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \quad 0 < \nu_j \leq 3; \quad 0 < w_{jk} \leq 1;$$

$$j, k = 1, \ldots, K; \quad t = 1, \ldots, T.$$

Here all coefficients, except the adjustment (error-correction) parameters, are set at their prior means, and the numbers in parentheses are their prior standard deviations. The r error-correction terms in each equation are now supposed known and unconstrained in a Bayesian specification framework, and thus included at full weight along with the constant term.

---

[9] In system (5.15a), the resulting ML estimator of Σ_u can be written as:

$$\tilde{\Sigma}_u = \left(\nabla \underline{Y} + \underline{A}_p(1) \underline{Y}_{-p}\right) \left(I_T - \nabla \underline{X}' (\nabla \underline{X} \nabla \underline{X}')^{-1} \nabla \underline{X}\right) \left(\nabla \underline{Y} + \underline{A}_p(1) \underline{Y}_{-p}\right)' / T,$$

where $\underset{(K \times T)}{\nabla \underline{Y}} = [\nabla \underline{Y}_1, \ldots, \nabla \underline{Y}_T]$, $\underset{(K \times T)}{\underline{Y}_{-p}} = [\underline{Y}_{1-p}, \ldots, \underline{Y}_{T-p}]$, $\underset{((1+K(p-1)) \times T)}{\nabla \underline{X}} = [\nabla \underline{X}_0, \ldots, \nabla \underline{X}_{T-1}]$ and

$$\underset{((1+K(p-1)) \times 1)}{\nabla \underline{X}_t} = \begin{bmatrix} 1 \\ \nabla \underline{Y}_t \\ \vdots \\ \nabla \underline{Y}_{t-p+2} \end{bmatrix}.$$

Also, as is traditional, the available prior information on the parameter vector of interest $\underline{\beta}_j^{\dagger}=(C_j, A_{j1,1}^{\dagger}, \ldots, A_{jj,1}^{\dagger}, \ldots, A_{jj,p-1}^{\dagger}, \ldots, A_{jK,p-1}^{\dagger})'$ can be written in the form

$$\underline{r}_j^{\dagger} = \underline{R}_j^{\dagger}\underline{\beta}_j^{\dagger} + \underline{v}_j^{\dagger} \quad \text{with } \underline{v}_j^{\dagger} \sim \text{NID}(0, I_{1+K(p-1)}); \qquad (5.18)$$
$$j = 1, \ldots, K,$$

where

$$\underset{((1+K(p-1))\times 1)}{\underline{r}_j^{\dagger}} = (0, 0, \ldots, 0, \ldots, 0, \ldots, 0)';$$

$$\underset{((1+K(p-1))\times(1+K(p-1)))}{\underline{R}_j^{\dagger}} = \text{diag}\left(0, \hat{\sigma}_1/\nu_j w_{j1}\hat{\sigma}_j, \ldots, 1/\nu_j, \ldots, (p-1)/\nu_j, \ldots, (p-1)\hat{\sigma}_K/\nu_j w_{jK}\hat{\sigma}_j\right);$$

$$\underset{((1+K(p-1))\times 1)}{\underline{\beta}_j^{\dagger}} = (C_j, A_{j1,1}^{\dagger}, \ldots, A_{jj,1}^{\dagger}, \ldots, A_{jj,p-1}^{\dagger}, \ldots, A_{jK,p-1}^{\dagger})'; \text{ and}$$

$$\underset{((1+K(p-1))\times 1)}{\underline{v}_j^{\dagger}} = (v_{j0}, v_{j1,1}^{\dagger}, \ldots, v_{jj,1}^{\dagger}, \ldots, v_{jj,p-1}^{\dagger}, \ldots, v_{jK,p-1}^{\dagger})'.$$

In this formulation, $\underline{r}_j^{\dagger}$ is a vector of zeros, $\underline{R}_j^{\dagger}$ a diagonal matrix with one zero on its diagonal corresponding to the constant term. Hence, $\underline{R}_j^{\dagger}\underline{R}_j^{\dagger\prime}$ is singular here as well, reflecting the improper flat prior on the constant term in each equation.

One important feature of this BVAR-EC model is that it allows the two sets of restrictions — the linear prior restrictions on $\underline{A}_j^{\dagger}$, $j=1,\ldots,p-1$, and the nonlinear cointegration restrictions on $A_p(1)$ — to be disentangled. Therefore, the conceptually distinct prior and reduced rank hypothesis implied by cointegration can each be investigated against facts. This model-building strategy may be useful for forecasting applications.

Having dealt with the determination of the prior and cointegration restrictions, we now turn to their incorporation in vector autoregressions. The proposed model, computed as part of the trial-and-error process, uses all currently available sample evidence to revise the prior probabilities, and the optimal shrinkage point estimators can thus be obtained via the use of Theil's mixed estimation technique. The validity of the final

forecasting model may also be judged internally by the important post-sample predictive testing, as was done in Litterman (1986a).

§5.3 Estimation of the Prespecified BVAR Model with Both Prior and Cointegration Restrictions

In general, the estimated values for a typical equation's parameter vector, say $\underset{\sim}{\beta}_j$ (or $\underset{\sim}{\beta}_j^{\dagger}$), can be presented solely in terms of a complete posterior density via Bayes' theorem, or alternatively, optimal point estimates through mixed estimation. For simplicity, however, we will focus on the point estimates of the regression coefficients using Theil's mixed estimation procedure.

Suppose a sample series of T observations is generated by the j-th equation of the system (5.10a). Stacking the T observations, we have

$$\underset{\sim}{Y}_j = \underset{\sim}{X}\underset{\sim}{\beta}_j + \underset{\sim}{u}_j; \qquad j = 1, \ldots, K, \tag{5.19}$$

where $\underset{\sim}{Y}_j = (Y_{j1}, \ldots, Y_{jT})'$; $(T\times 1)$

$$\underset{\sim}{X} = (\underset{\sim}{X}_0, \ldots, \underset{\sim}{X}_{T-1})' \text{ with } \underset{\sim}{X}_t = \begin{bmatrix} 1 \\ Y_{1,t} \\ \vdots \\ Y_{1,t-p+1} \\ \vdots \\ Y_{K,t} \\ \vdots \\ Y_{K,t-p+1} \end{bmatrix}; \quad (T\times(1+Kp)) \quad ((1+Kp)\times 1)$$

$$\underset{\sim}{\beta}_j = (C_j, A_{j1,1}, \cdots, A_{j1,p}, \cdots, A_{jK,1}, \cdots, A_{jK,p})'; \text{ and} \quad ((1+Kp)\times 1)$$

$$\underset{\sim}{u}_j = (u_{j1}, \ldots, u_{jT})' \sim NID(0, \sigma_{jj}^2 I_T). \quad (T\times 1)$$

Then, if we employ this type of sample information in conjunction with the prior information contained in the process (5.14)

$$\underset{\sim}{r}_j = \underset{\sim}{R}_j\underset{\sim}{\beta}_j + \underset{\sim}{v}_j; \qquad j = 1, \ldots, K,$$

where $\underline{v}_j \sim NID(0,\ I_{1+Kp})$, and $E(\underline{u}_j\underline{v}_j') = 0$, the shrinkage estimator of $\underline{\beta}_j$ is:

$$\hat{\underline{\beta}}_j = [\underline{R}_j'\underline{R}_j + \sigma_{jj}^{-2}\underline{X}'\underline{X}]^{-1}[\underline{R}_j'\underline{r}_j + \sigma_{jj}^{-2}\underline{X}'\underline{Y}_j]. \qquad (5.20)$$

Now suppose the observed data chosen is produced by the set of the equations of the system (5.15a)

$$\nabla\underline{Z}_j = \nabla\underline{X}\underline{\beta}_j^{\dagger} + \underline{u}_j; \qquad j = 1, \ldots, K, \qquad (5.21)$$

where $\nabla\underline{Z}_j = \underset{(T\times 1)}{\begin{bmatrix} \nabla Y_{j1} + \hat{\omega}_{j1}\hat{\varepsilon}_{1,1-p} + \ldots + \hat{\omega}_{jr}\hat{\varepsilon}_{r,1-p} \\ \vdots \\ \nabla Y_{jT} + \hat{\omega}_{j1}\hat{\varepsilon}_{1,T-p} + \ldots + \hat{\omega}_{jr}\hat{\varepsilon}_{r,T-p} \end{bmatrix}}$

$$= \underset{(T\times 1)}{\begin{bmatrix} \nabla Y_{j1} \\ \vdots \\ \nabla Y_{jT} \end{bmatrix}} + \left(\underset{(1\times r)}{\begin{bmatrix} \hat{\omega}_{j1} & \cdots & \hat{\omega}_{jr} \end{bmatrix}} \underset{(r\times T)}{\begin{bmatrix} \hat{\varepsilon}_{1,1-p} & \cdots & \hat{\varepsilon}_{1,T-p} \\ \vdots & & \vdots \\ \hat{\varepsilon}_{r,1-p} & \cdots & \hat{\varepsilon}_{r,T-p} \end{bmatrix}} \right)'$$

$$= \underset{(T\times 1)}{\begin{bmatrix} \nabla Y_{j1} \\ \vdots \\ \nabla Y_{jT} \end{bmatrix}} + \left(\underset{(1\times r)}{\begin{bmatrix} \hat{\omega}_{j1} & \cdots & \hat{\omega}_{jr} \end{bmatrix}} \underset{(r\times K)}{\begin{bmatrix} \hat{\alpha}_{11} & \cdots & \hat{\alpha}_{1K} \\ & \vdots & \\ \hat{\alpha}_{r1} & \cdots & \hat{\alpha}_{rK} \end{bmatrix}} \underset{(K\times T)}{\begin{bmatrix} Y_{1,1-p} & \cdots & Y_{1,T-p} \\ \vdots & & \vdots \\ Y_{K,1-p} & \cdots & Y_{K,T-p} \end{bmatrix}} \right)';$$

$$\nabla\underline{X} = \underset{(T\times(1+K(p-1)))}{(\nabla\underline{X}_0, \ldots, \nabla\underline{X}_{T-1})'} \text{ with } \nabla\underline{X}_t = \underset{((1+K(p-1))\times 1)}{\begin{bmatrix} 1 \\ \nabla Y_{1,t} \\ \vdots \\ \nabla Y_{1,t-p+2} \\ \vdots \\ \nabla Y_{K,t} \\ \vdots \\ \nabla Y_{K,t-p+2} \end{bmatrix}};$$

$$\underline{\beta}_j^{\dagger} = \underset{((1+K(p-1))\times 1)}{(C_j, A_{j1,1}^{\dagger}, \ldots, A_{j1,p-1}^{\dagger}, \ldots, A_{jK,1}^{\dagger}, \ldots, A_{jK,p-1}^{\dagger})'}; \text{ and}$$

$$\underline{u}_j = \underset{(T\times 1)}{(u_{j1}, \ldots, u_{jT})'} \sim NID(0,\ \sigma_{jj}^2 I_T).$$

Then, provided that all the error-correction terms are known before estimation, by combining the data above with the prior held in the

process (5.18)

$$\underline{r}_j^{\dagger} = \underline{R}_j^{\dagger}\underline{\beta}_j^{\dagger} + \underline{v}_j^{\dagger}; \qquad j = 1, \ldots, K,$$

where $\underline{v}_j^{\dagger} \sim NID(0, I_{1+K(p-1)})$, and $E(\underline{u}_j\underline{v}_j^{\dagger\prime}) = 0$,

we obtain the mixed estimator for $\underline{\beta}_j^{\dagger}$ of the j-th equation

$$\begin{aligned}\hat{\underline{\beta}}_j^{\dagger} &= [\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{X}]^{-1}[\underline{R}_j^{\dagger\prime}\underline{r}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{Z}_j] \\ &= [\sigma_{jj}^{2}\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger} + \nabla\underline{X}'\nabla\underline{X}]^{-1}\nabla\underline{X}'\nabla\underline{Z}_j, \quad \text{as } \underline{r}_j^{\dagger} \triangleq 0. \end{aligned} \tag{5.22}$$

This estimator looks akin to the LS estimator in first differences except for $\sigma_{jj}^{2}\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger}$. What is more, like the estimator given by (5.20), it is constructed for each of the K equations of the system separately and thus can be easily applied.

When the coefficients of a Bayes set-up have been estimated in this way from the available prior and data, the estimated process may be used for prediction and simulation.

§5.4 Prediction of the Mixed-Estimated BVAR Model

Similar principles to the Bayesian prediction procedure outlined in Chapter 3 can be applied here to establish predictive densities, in the form of a multivariate-t distribution, for both cases. In particular, if point predictions are desired, the optimal out-of-sample, multi-step predictor, based on the minimization of squared prediction error, can be derived quite simply from the shrinkage estimator or the mixed estimator $\hat{\underline{\beta}}_j$ (or $\hat{\underline{\beta}}_j^{\dagger}$) period by period.

Suppose we wish to derive a set of point predictions for h future values on the j-th variable Y_j, say $\tilde{\underline{Y}}_j$, which is assumed to satisfy the same regression model generating its sample observations $\underline{Y}_j$ as in (5.19); i.e.

$$\tilde{\underline{Y}}_j = \tilde{\underline{X}}\underline{\beta}_j + \tilde{\underline{u}}_j; \qquad j = 1, \ldots, K, \tag{5.23}$$

where $\underset{(h\times 1)}{\tilde{\underline{Y}}_j} = (Y_{j,T+1}, \ldots, Y_{j,T+h})'$;[10]

$$\underset{(h\times(1+Kp))}{\tilde{\underline{X}}} = (\underline{X}_T, \ldots, \underline{X}_{T+h-1})' \text{ with } \underline{X}_t = \underset{((1+Kp)\times 1)}{\begin{bmatrix} 1 \\ Y_{1,t} \\ \vdots \\ Y_{1,t-p+1} \\ \vdots \\ Y_{K,t} \\ \vdots \\ Y_{K,t-p+1} \end{bmatrix}};$$

$$\underset{((1+Kp)\times 1)}{\underline{\beta}_j} = (C_j, A_{j1,1}, \ldots, A_{j1,p}, \ldots, A_{jK,1}, \ldots, A_{jK,p})'; \text{ and}$$

$$\underset{(h\times 1)}{\tilde{\underline{u}}_j} = (u_{j,T+1}, \ldots, u_{j,T+h})' \sim \text{NID}(0, \sigma_{jj}^2 I_h) \text{ with } E(\underline{u}_j \tilde{\underline{u}}_j') = 0.$$

Then, given σ_{jj} and forecast values in $\tilde{\underline{X}}$ as if they were actual observed data as the date is advanced into the future, the optimal point predictions of Y_j over the h future periods would be

$$\begin{aligned} \hat{\underline{Y}}_j &= E(\tilde{\underline{Y}}_j) = \tilde{\underline{X}}\hat{\underline{\beta}}_j \\ &= \tilde{\underline{X}}[\underline{R}_j'\underline{R}_j + \sigma_{jj}^{-2}\underline{X}'\underline{X}]^{-1}[\underline{R}_j'\underline{r}_j + \sigma_{jj}^{-2}\underline{X}'\underline{Y}_j]; \qquad (5.24) \\ &\quad j = 1, \ldots, K. \end{aligned}$$

The associated covariance matrix is given by

$$\begin{aligned} &E(\tilde{\underline{Y}}_j - E(\tilde{\underline{Y}}_j))(\tilde{\underline{Y}}_j - E(\tilde{\underline{Y}}_j))' \\ &= E(\tilde{\underline{X}}\underline{\beta}_j + \tilde{\underline{u}}_j - \tilde{\underline{X}}\hat{\underline{\beta}}_j)(\tilde{\underline{X}}\underline{\beta}_j + \tilde{\underline{u}}_j - \tilde{\underline{X}}\hat{\underline{\beta}}_j)' \\ &= E(\tilde{\underline{X}}(\underline{\beta}_j - \hat{\underline{\beta}}_j) + \tilde{\underline{u}}_j)(\tilde{\underline{X}}(\underline{\beta}_j - \hat{\underline{\beta}}_j) + \tilde{\underline{u}}_j)' \\ &= \tilde{\underline{X}}E[(\underline{\beta}_j - \hat{\underline{\beta}}_j)(\underline{\beta}_j - \hat{\underline{\beta}}_j)']\tilde{\underline{X}}' + E(\tilde{\underline{u}}_j\tilde{\underline{u}}_j') \\ &= \tilde{\underline{X}}\bar{\underline{V}}_j\tilde{\underline{X}}' + \sigma_{jj}^2 I_h \\ &= \tilde{\underline{X}}(\underline{R}_j'\underline{R}_j + \sigma_{jj}^{-2}\underline{X}'\underline{X})^{-1}\tilde{\underline{X}}' + \sigma_{jj}^2 I_h. \qquad (5.25) \end{aligned}$$

Suppose, on the other hand, we wish to derive predictions for Y_j in the next h periods through the alternative regression model (5.21);

[10] Here, the "hat" over the predicted value is omitted to simplify the presentation.

viz.

$$\nabla\underline{\tilde{Y}}_j = \nabla\underline{\tilde{X}}\underline{\beta}_j^{\dagger} - \hat{\omega}_{j1}\underline{\tilde{\varepsilon}}_1 - \ldots - \hat{\omega}_{jr}\underline{\tilde{\varepsilon}}_r + \underline{\tilde{u}}_j$$
$$= \nabla\underline{\tilde{X}}\underline{\beta}_j^{\dagger} - (\underline{\hat{\omega}}_j \otimes I_h)\underline{\tilde{\varepsilon}} + \underline{\tilde{u}}_j; \qquad j = 1, \ldots, K, \tag{5.26}$$

where $\underset{(h\times 1)}{\nabla\underline{\tilde{Y}}_j} = (\nabla\underline{Y}_{j,T+1}, \ldots, \nabla\underline{Y}_{j,T+h})'$;

$$\underset{(h\times(1+K(p-1)))}{\nabla\underline{\tilde{X}}} = (\nabla\underline{X}_T, \ldots, \nabla\underline{X}_{T+h-1})' \text{ with } \nabla\underline{X}_t = \underset{((1+K(p-1))\times 1)}{\begin{bmatrix} 1 \\ \nabla Y_{1,t} \\ \vdots \\ \nabla Y_{1,t-p+2} \\ \vdots \\ \nabla Y_{K,t} \\ \vdots \\ \nabla Y_{K,t-p+2} \end{bmatrix}};$$

$$\underline{\beta}_j^{\dagger} = \underset{((1+K(p-1))\times 1)}{(C_j, A_{j1,1}^{\dagger}, \ldots, A_{j1,p-1}^{\dagger}, \ldots, A_{jK,1}^{\dagger}, \ldots, A_{jK,p-1}^{\dagger})'};$$

$$\underset{(1\times r)}{\underline{\hat{\omega}}_j} = (\hat{\omega}_{j1}, \ldots, \hat{\omega}_{jr});$$

$$\underset{(hr\times 1)}{\underline{\tilde{\varepsilon}}} = (\underline{\tilde{\varepsilon}}_1', \ldots, \underline{\tilde{\varepsilon}}_r')' \text{ with } \underline{\tilde{\varepsilon}}_i = \underset{(h\times 1)}{\begin{bmatrix} \hat{\underline{\varepsilon}}_{i,T+1-p} \\ \vdots \\ \hat{\underline{\varepsilon}}_{i,T+h-p} \end{bmatrix}}; \text{ and}$$

$$\underset{(h\times 1)}{\underline{\tilde{u}}_j} = (u_{j,T+1}, \ldots, u_{j,T+h})' \sim \text{NID}(0, \sigma_{jj}^2 I_h) \text{ with } E(\underline{u}_j\underline{\tilde{u}}_j') = 0.$$

Then, given $\nabla\underline{\tilde{X}}$, σ_{jj}, $\underline{\hat{\omega}}_j$ and $\underline{\tilde{\varepsilon}}$, prior to making a forecast, the optimal point predictions of the differenced series ∇Y_j would be

$$\nabla\underline{\hat{Y}}_j = E(\nabla\underline{\tilde{Y}}_j)$$
$$= \nabla\underline{\tilde{X}}\underline{\hat{\beta}}_j^{\dagger} - (\underline{\hat{\omega}}_j \otimes I_h)\underline{\tilde{\varepsilon}}$$
$$= \nabla\underline{\tilde{X}}[\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{X}]^{-1}[\underline{R}_j^{\dagger\prime}\underline{r}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{Z}_j] - (\underline{\hat{\omega}}_j \otimes I_h)\underline{\tilde{\varepsilon}}$$
$$= \nabla\underline{\tilde{X}}[\sigma_{jj}^2\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger} + \nabla\underline{X}'\nabla\underline{X}]^{-1}\nabla\underline{X}'\nabla\underline{Z}_j - (\underline{\hat{\omega}}_j \otimes I_h)\underline{\tilde{\varepsilon}}; \tag{5.27}$$
$$j = 1, \ldots, K; \quad \text{as } \underline{r}_j^{\dagger} = 0,$$

with the covariance matrix

$$E(\nabla\underline{\tilde{Y}}_j - E(\nabla\underline{\tilde{Y}}_j))(\nabla\underline{\tilde{Y}}_j - E(\nabla\underline{\tilde{Y}}_j))'$$
$$= E(\nabla\underline{\tilde{X}}(\underline{\beta}_j^{\dagger} - \underline{\hat{\beta}}_j^{\dagger}) + \underline{\tilde{u}}_j)(\nabla\underline{\tilde{X}}(\underline{\beta}_j^{\dagger} - \underline{\hat{\beta}}_j^{\dagger}) + \underline{\tilde{u}}_j)'$$

$$= \nabla\tilde{\underline{X}}\overline{\underline{V}}_j^{\dagger}\nabla\tilde{\underline{X}}' + E(\tilde{\underline{u}}_j\tilde{\underline{u}}_j')$$
$$= \nabla\tilde{\underline{X}}(\underline{R}_j^{\dagger}\underline{R}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{X})^{-1}\nabla\tilde{\underline{X}}' + \sigma_{jj}^{2}I_h. \qquad (5.28)$$

Moreover, since $Y_{jt}=\nabla Y_{jt}+Y_{j,t-1}$, the optimal point predictions of h future values for the original level series Y_j can be derived recursively from (5.27) by using the chain rule of forecasting or backward substitution. This takes the calculated one-step-ahead forecast as the basis for a two-step-ahead forecast and so on, until the h-step-ahead forecast has been reached.[11] We thus have

$$\hat{Y}_{j,T+1} = \nabla\hat{Y}_{j,T+1} + Y_{j,T},$$
$$\hat{Y}_{j,T+2} = \nabla\hat{Y}_{j,T+2} + \hat{Y}_{j,T+1} = \nabla\hat{Y}_{j,T+2} + \nabla\hat{Y}_{j,T+1} + Y_{j,T},$$
$$\cdots \quad \cdots$$
$$\hat{Y}_{j,T+h} = \nabla\hat{Y}_{j,T+h} + \cdots + \nabla\hat{Y}_{j,T+1} + Y_{j,T} = Y_{j,T} + \sum_{i=0}^{h-1}\nabla\hat{Y}_{j,T+h-i}.$$

In matrix form, we find that

$$\hat{\underline{Y}}_j^{\dagger} = \underline{Y}_{jT} + i_h\nabla\hat{\underline{Y}}_j$$
$$= \underline{Y}_{jT} + i_h\nabla\tilde{\underline{X}}[\sigma_{jj}^{2}\underline{R}_j^{\dagger}\underline{R}_j^{\dagger} + \nabla\underline{X}'\nabla\underline{X}]^{-1}\nabla\underline{X}'\nabla\underline{Z}_j \qquad (5.29)$$
$$- i_h(\hat{\underline{\omega}}_j\otimes I_h)\tilde{\underline{\varepsilon}}; \qquad j = 1, \ldots, K,$$

where $\hat{\underline{Y}}_j^{\dagger} = (\hat{Y}_{j,T+1}, \ldots, \hat{Y}_{j,T+h})'$ $(h\times 1)$;

$\underline{Y}_{jT} = (Y_{j,T}, \ldots, Y_{j,T})'$ $(h\times 1)$;

$$i_h = \begin{bmatrix} 1 & & 0 \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{bmatrix}; \quad (h\times h)$$

and

$$\nabla\hat{\underline{Y}}_j = (\nabla\hat{Y}_{j,T+1}, \ldots, \nabla\hat{Y}_{j,T+h})', \quad (h\times 1)$$

with the covariance matrix given by

$$E(\tilde{\underline{Y}}_j^{\dagger} - E(\tilde{\underline{Y}}_j^{\dagger}))(\tilde{\underline{Y}}_j^{\dagger} - E(\tilde{\underline{Y}}_j^{\dagger}))'$$

[11] See Pindyck & Rubinfeld (1991), Chapter 18, for further details.

$$
\begin{aligned}
&= E\big(\underline{Y}_{jT} + i_h\nabla\tilde{\underline{Y}}_j - (\underline{Y}_{jT} + i_h\nabla\hat{\underline{Y}}_j)\big) \\
&\qquad \big(\underline{Y}_{jT} + i_h\nabla\tilde{\underline{Y}}_j - (\underline{Y}_{jT} + i_h\nabla\hat{\underline{Y}}_j)\big)' \\
&= E\big(i_h(\nabla\tilde{\underline{Y}}_j - \nabla\hat{\underline{Y}}_j)\big)\big(i_h(\nabla\tilde{\underline{Y}}_j - \nabla\hat{\underline{Y}}_j)\big)' \\
&= i_h E\big[(\nabla\tilde{\underline{Y}}_j - \nabla\hat{\underline{Y}}_j)(\nabla\tilde{\underline{Y}}_j - \nabla\hat{\underline{Y}}_j)'\big] i_h' \\
&= (i_h\nabla\tilde{\underline{X}})\big[\underline{R}_j^{\dagger\prime}\underline{R}_j^{\dagger} + \sigma_{jj}^{-2}\nabla\underline{X}'\nabla\underline{X}\big]^{-1}(i_h\nabla\tilde{\underline{X}})' + i_h\sigma_{jj}^{2}I_h i_h'. \qquad (5.30)
\end{aligned}
$$

A comparison of Eq.(5.25) with Eq.(5.30) should make it clear that only in the second case can the valid reduced-rank restrictions be explicitly incorporated, along with the prior restrictions, to improve forecast accuracy over long forecast horizons. The two kinds of optimal point predictions (or mixed estimator forecasts) will be calculated in turn in a multicountry context, and then analysed and compared systematically in subsequent forecasting applications.

In addition, the h×1 mean vector $\hat{\underline{Y}}_j$ or $\hat{\underline{Y}}_j^{\dagger}$, as well as the squared roots of the diagonal elements of the corresponding h×h covariance matrix (5.25) or (5.30) can be used to construct confidence intervals at a required critical value for the h elements of $\tilde{\underline{Y}}_j$ or $\tilde{\underline{Y}}_j^{\dagger}$, respectively. It is now generally recognised that I(1) variables can only be forecast with increasingly wide confidence intervals, whereas stationary, cointegrated linear combinations of such variables have finite confidence intervals as the forecast horizon lengthens [Clements & Hendry (1995), p.127].

§5.5 Conclusion

In this chapter, we have discussed the construction, estimation and prediction of both BVAR and BVAR-EC models for integrated and cointegrated variables. Insofar as macroeconometric models forecast the future, a traditional BVAR formulation utilizes the prior restrictions to increase the model's forecasting capability for small sample sizes. A statistical EC representation, in contrast, capitalizes on the long-run constraints implied by cointegration to enhance the model's predictive power over long forecast horizons.

The focus of this study is thus on whether prior and cointegration restrictions can be both imposed explicitly, rather than implicitly, during the estimation stage. This point has hitherto received little, if any, attention in the literature but will most probably have significant implications for the practice of BVAR forecasting.

Systematic exploration of the sensitivity of forecasts to decisions taken concerning the priors on parameter values gives no guarantee that we can obtain the most suitable prior for the period to be forecast. Litterman's vector random walk prior, as well as its possible linear transformations, is suitable for most economic variables, but not all. As is noted in Artis *et al.* (1990a, p.350):

> *Our experience has been that the gains from adopting the Bayesian approach may be offset by setting up an inappropriate prior.*

To rephrase this, the order of precedence in terms of forecast accuracy is that a good Bayesian will beat a non-Bayesian, which will in turn beat a poor Bayesian [Granger (ed) (1991), p.18].

CHAPTER 6

ESTIMATION AND HYPOTHESIS TESTING IN MULTIVARIATE SYSTEMS OF EQUATIONS

§6.1 Introduction

Given the modelling methodology formalized in the preceding chapters, this and next chapters move directly through estimation, testing and construction of appropriate BVAR forecasting models. They present empirical evidence, testing results and forecast accuracy assessment for both multivariate single- and cross-country cases. The main goal in this part of the study is to explore the properties of and relationships between the time series involved; to investigate the implications of imposing prior and long-run constraints for forecast accuracy at small sample sizes; and to establish a general quarterly BVAR model within a multicountry context. A practical six-step model-building procedure is employed to this purpose.

(1) Pick the four major national economies in Europe, i.e.: the United Kingdom, Germany, France and Italy.

(2) Select a set of distinct aggregate variables of interest for use in the analysis:

(a) choose a similar set of macroeconomic aggregates: namely, the real (inflation-adjusted) GDP/GNP, Money Supply (as measured by M0/M1), Consumer Price Index (1990/1991=100, CPI for short), and (possibly) 3-Month Treasury Bill Rate (TBR) for one country after another, assuming that each member economy is described by these macroeconomic indicators;[1]

[1] The nature of the intertemporal correlations among these variables within the context of VAR processes has received considerable attention in the empirical business cycle literature, e.g., Blanchard (1989), Gali (1992), Litterman & Weiss (1985), Sims (1980a), and Stock & Watson (1988a, 1988b).

(b) take natural logs of all series (except Treasury bills); and

(c) plot these series both individually and jointly, and then analyse them.

(3) Examine the presence of long-run cointegrating relationships:

(a) test for unit roots or the degree of integration in each individual series;

(b) test for the maximum lag length of the VAR model, p, for those variables that are to be modelled;

(c) test for the rank order of the cointegration space, r, for those integrated variables appearing in the model, provided that the value of p is specified;

(d) estimate cointegrating vectors (CVs) $\underline{\alpha}$, adjustment matrix $\underline{\omega}$ and long-run matrix $\underline{A}_p(1)\left(=\underline{\omega}\underline{\alpha}'\right)$, and save the residual vectors $\hat{\underline{\varepsilon}}_t\left(=\hat{\underline{\alpha}}'\underline{Y}_t\right)$ associated with the r CVs, given that the value of r is fixed; and

(e) graph the saved residuals of the CV(s) to check for possible nonstationary patterns, and/or test them to see if they are I(0).

(4) Reveal the best scalar settings of the hyperparameters in both symmetric and general priors over the whole simulated 'out-of-sample' forecast period:

(a) run a system of unrestricted univariate OLS models to get benchmark Theil U's;

(b) run a system of restricted univariate models with a standard value for the overall tightness parameter ν_j. If the Theil U's are getting worse, loosen up on the diagonal elements; and

(c) run a standard symmetric prior for the tightness parameters on the 'other' variables in the relative tightness function f(i,j), w_{jk}. If the Theil U's are getting worse, tighten up on the off-diagonal elements.

(5) Formulate the empirical BVAR and BVAR-EC forecasting models:

(a) set up a BVAR model for the chosen set of variables in levels with the prespecified stochastic prior restrictions imposed on its lagged coefficients;

(b) set up a BVAR-EC model in (first) differences if there is cointegration between the integrated (I(1)) variables:

(i) fit the r lagged estimated cointegrating residuals $\hat{\underline{\varepsilon}}_{t-p}(=\hat{\underline{\alpha}}'\underline{Y}_{t-p})$ into each of the recursive equations of the bivariate and/or multivariate error-correction models (ECMs) as additional regressor(s);

(ii) impose the prior distributions associated with the 'best' setting of the hyperparameters on all coefficients except for the constant and error-correction terms. Then estimate the resulting equations (subject to both prior and cointegration restrictions) jointly using the mixed rolling estimation method;

(iii) use identities $Y_{jt} \equiv Y_{j,t-1} + \nabla Y_{jt}$, $j=1,\ldots,K$, to transform the successive forecasted difference series into level series; and

(iv) update the r lagged error-correction terms $(\hat{\underline{\varepsilon}}_{t-p})$ at each forecast step, assuming that the 'equilibrium relationships' defined by the r CVs $\hat{\underline{\alpha}}$ will continue to hold over the forecasting horizons.

(6) Compute a series of ex-ante quarterly forecasts based on the estimated BVAR and BVAR-EC models and evaluate the relative forecasting performance over time:

(a) use BVAR and BVAR-EC models to generate a series of 1-step and multi-step quarterly (point) forecasts for the leading European industrial economies. Optimally, a Kalman filter (KF) estimation method should be used to facilitate the remaining successive forecasts by re-estimating coefficients in the models prior to each forecast period. Afterwards, the anti-log transformation of the forecasted series could be taken; then

(b) test efficiency and unbiasedness properties of the quarterly forecasts and undertake a systematic comparison of the forecasts across alternative models.

The systematic six-step procedure offered here is designed to be flexible and readily implemented in building a BVAR or a BVAR-EC model, and will be followed, step by step, in the sections that follow.

§6.2 Mainstream Data-Mining (MDM) — a Suitable Way of Database Assembly from Various Currently Existing Databanks

The term 'data mining', or sometimes 'specification searches', means the re-use of the same limited data both to estimate and to revise a model, since a model must be revised in the light of data evidence unless it emerges perfect at the first attempt. The industrious implication of the word 'mining' suggests that the activity may be productive. However, a better approach in the process of model construction is the explicit use of prior and cointegration restrictions, as remarked in the previous chapters. For further elaboration and references see Leamer (1991, pp.235-58). It has been noticed that in practical time series analyses, a quarterly model is often used for the study of important phenomena like business cycles [on this see Bodkin, Klein & Marwah (1991), pp.89-90]. But sometimes, quarterly models are constrained by the availability of data. To create an adequate quarterly data set upon which to base analyses is therefore an important aspect of macroeconometric work.

6.2.1 Accumulation, Selection and Transformation of Raw Data Series

In this section, we start with the selection of the raw data on a quarterly basis, and then move on to the common instantaneous transformations of the data series. Usually, collecting a large consistent data set (for a reasonable span of time) requires access to the various existing large and high-quality databanks and networked information servers. Such a database is held on Manchester Information Datasets and Associated Services (MIDAS), which provides a National Datasets Service to the UK academic community.

6.2.1a *The Selection of the Data*

All aggregate quarterly time-series data are obtained from a database first assembled from the ONS Macroeconomic Time Series (MTS), the International Monetary Fund (IMF) International Financial Statistics (IFS), and the OECD Main Economic Indicators (MEI)

databanks. These can be accessed freely via MIDAS, and fall into two groups. The data for the variables labelled CPI and TBR are seasonally unadjusted and chosen from the IMF IFS sources, while the data for the real GDP/GNP and MO/M1 are seasonally adjusted and taken to be the OECD MEI and ONS MTS data.[2] The series are subsequently reset so that all are based upon the current benchmark year or updated using recent monthly issues of the relevant publications. Further details, including sources, names and explanations, of the data chosen for the major four (and other) European economies are given in Appendix C of this thesis.

6.2.1b *The Transformation of the Data*

When working with potentially nonstationary processes, certain transformations of the original series to render it I(0) by way of taking differences, log-differences and establishing cointegrating combinations are often required. There are two main advantages of taking logs instead of levels. One is that taking natural logs of exponentially growing series helps reduce heteroskedasticity of the empirical series, as the series with a fixed percentage rise will be linearized. Although taking logs cannot remove changes in the variance *per se*, the assumption of constant variance may prove to be a more reasonable approximation for a logarithmic series. Examples would be real GDP/GNP, monetary aggregates or consumer prices. Non-trending series, such as interest rates, should be left in levels. The second is that the usual linear relationships among the absolute changes of the candidates being modelled become more meaningful log-linear channels that convey relative changes between one of the variables and the past of the other explanatory variables involved. The sign and size of the additive error corrected for each of the dependent variables in the error correction mechanism will depend on the direction and magnitude of the proportionate or percentage change of the variable considered each period. These

[2]Many released records accessed by and downloaded from terminals in a network are 'long consistent', but one should be aware of possible limitations of almost any database to be used at this stage. Such limitations may stem either from changes in economic structure and policy regimes or from changes in definition of a given series and delays of new releases.

points have important implications in empirical applications.

Here, all series but short-term interest rates are first transformed to logs prior to analysis. The macro data used with each country will be a group of four as follows:

LGDP/LGNP = log of GDP/GNP, at constant prices (adj.)
LM0/LM1 = log of Money Supply M0/M1 (adj.)
LCPI = log of the Consumer Price Index (1990/1991=100) (orig.)
TBR = 3-Month Treasury Bill Rate (orig.).

Whilst practice might vary on this, this four-candidate (K=4) quarterly system is considered for three reasons: (a) the forecasts of these macroeconomic variables are commonly found in public discussion [cf. Artis *et al.* (1990a), Litterman & Weiss (1985), Sims (1980a), and Wallis (1989)]; (b) the choice of this subset has been found to be adequate for the purpose of conducting a real-time forecast comparison between the BVARs and other conventional competing techniques commercially available [cf. Artis *et al.* (1990a, 1990b), and Spencer (1993)]; and (c) the official macroeconomic data samples on the subset can be easily assembled, updated and transformed in ASCII files from ONS, IMF and OECD sources.

In this work, the sample period for the UK is 1969Q2-1996Q3 inclusive; for Germany 1975Q3-1992Q4; for France 1979Q1-1996Q3; and for Italy 1977Q2-1996Q3, the latest available at the time of the analysis. Hence, the precise estimation periods used do vary with specific country cases. It can be seen from the data appendix that, in almost all of these countries, yield on T-bills appears to be the shortest one among existing series. Thus, if it is included, it restricts the length of the sample. Moreover, in practical model building, all the associated tests will be based on a given subset of data through 1990Q4, reserving at least 2-years data, or 8 data points, from 1991Q1 onwards for simulated 'out-of-sample' multi-period forecasts.

However, before testing and finally specifying a 'good' forecasting model based on the available prior and sample information, a preliminary graphical inspection of the level series, after the

necessary transformations, can be a useful adjunct to formal models. See Clements and Hendry (1992a, p.9). After all, forecasters should know where the economy is before providing a reasonable picture of where it is going. See Pain and Britton (1992, p.87) among others.

6.2.2 Plotting, Survey and Analysis of the Available Data Samples

The observed series are plotted against time, both individually and jointly, to establish what are the dominant features of the series that a model will need to capture. These graphics may well indicate whether the series is subject to structural changes, either throughout the series, or at particular points in time. They may also suggest whether a group of time-series trajectories move roughly together or bring to light some unexpected effects. In fact, it is the examination of the historical data, combined with prior knowledge of the nature of the series, which provides the basis for model specification.

Presented below are the time plots of the four transformed macroeconomic indicators for the four major industrial countries in Europe over both the 'inside-sample' and 'outside-sample' periods. All these graphs are depicted through use of GiveWin under Windows and examined especially within the sample. Over that period, five important episodes are of interest: the rapid growth in overall activity in 1973 and 1987-88 (with the subsequent inflationary pressures), and the recessions of 1974-5, 1979-81 and 1991-2.

Real GDP/GNP growth

Figs. 6.1 and 6.2 show a block of four time plots for British LGDP (1990 Prices), German LGNP (1985 Prices), French LGDP (1980 Prices) and Italian LGDP (1990 Prices) at seasonally adjusted annual rates, and also a joint one of these four, over the period 1971Q1-1992Q4. There is a clear long-run upward trend in all four series. But, in addition, the plots of these series show marked close cyclical behaviour across different countries during the period of fit as the economies move from boom to recession and back again. Indeed, we

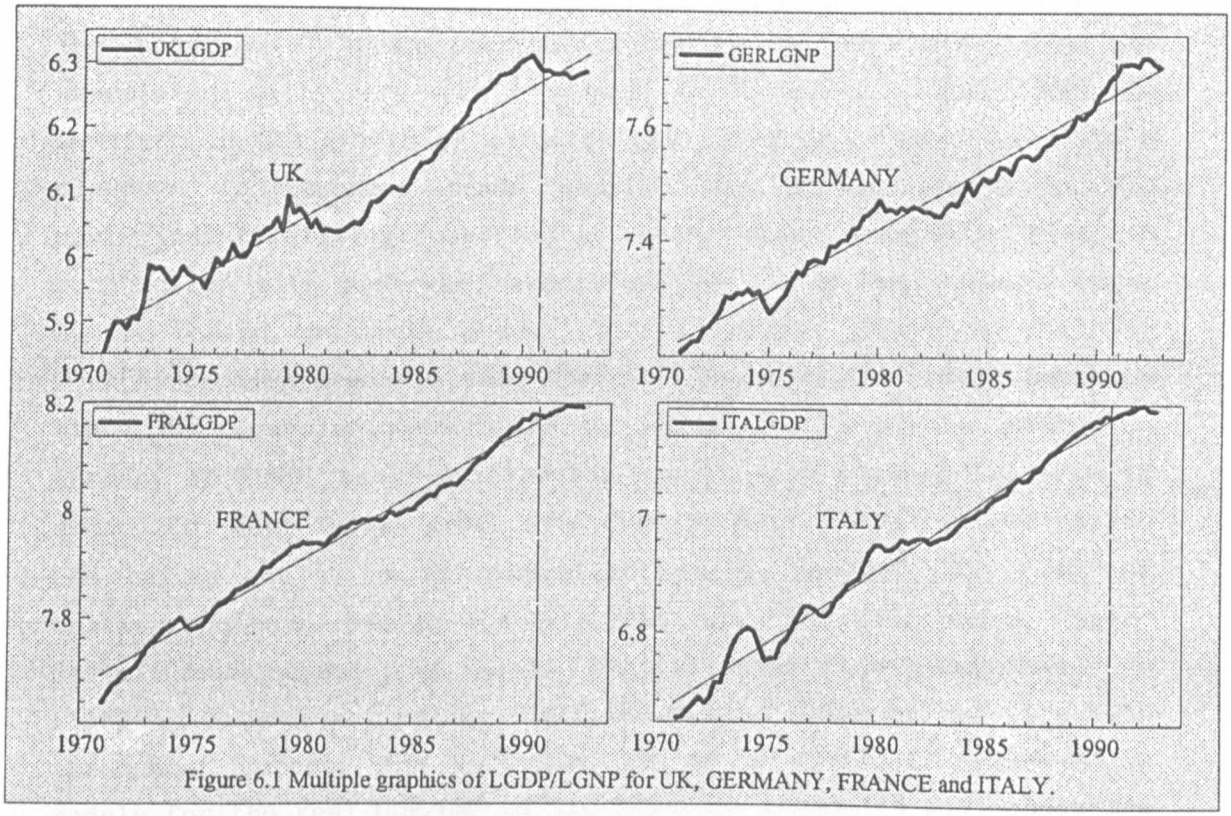

Figure 6.1 Multiple graphics of LGDP/LGNP for UK, GERMANY, FRANCE and ITALY.

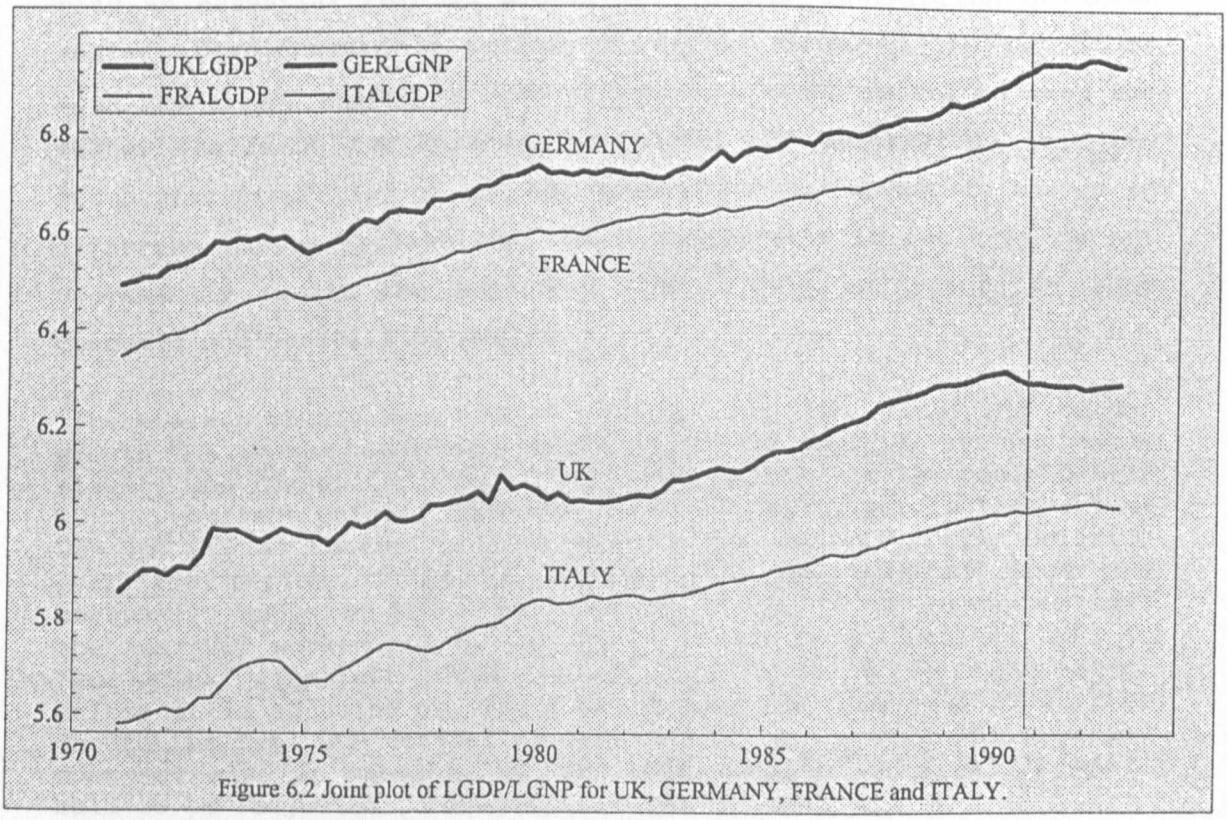

Figure 6.2 Joint plot of LGDP/LGNP for UK, GERMANY, FRANCE and ITALY.

would probably have ascertained this from the economic history of the period, without even looking at the graphs.

The boom in European economic activity in the late 1980s was associated with very rapid growth of European trade, and the downturn in the growth of activity was inevitably associated with a slowdown in European trade growth. The strong growth of the newly-industrialising countries in the Far East along with China, in conjunction with post-war reconstruction in the Gulf, was a major factor behind the trade growth. Within Europe, growth in the 'Big Four' economies began to slow down at the end of 1989, but the early signs of recession were muted by exceptionally strong growth in Germany in 1990, as re-unification resulted in a sharp increase in demand for West German goods from East Germany. German reunification has turned out to be an expensive process, however. The costs of unification have created a substantial and continuing public sector deficit and caused the German central bank (the Bundesbank) to tighten its monetary policy. This acted as a break on output growth throughout Europe. The incorporation of a cyclical component in models for the real GDP/GNP of the European economies will therefore play an essential role in providing a better description of these series. Developments in Germany have also extended the period of slow growth and helped to make growth more synchronised between the member states of the European Community. The observation that the stochastic properties of the aggregates examined appear to change in different phases implies that their characteristic features do not necessarily remain the same over time.[3] This is typical of many economic and social time series.

[3]There are probably some other reasons for the lingering output growth in these countries, particularly in the UK at her early stages. One reason for this is that what we want is the real GDP (or GDP at constant prices). But what we can actually get is the nominal GDP (or GDP at current prices) and the GDP Implicit Price Deflator over the period 1971Q4 to 1992Q4. Although both of them are seasonally adjusted and rise smoothly, there is no guarantee that the real GDP worked out by the ratio of the first (nominal GDP) to the second (implicit price deflator) will go up in every step, unless it is adjusted on itself as we will see from post-1993. Even though the real GDP derived in this way (from the OECD MEI sources) does not rise as smoothly as the usual adjusted data, it is still said to be 'adjusted', because it is not original.

Money supply

Recently, most European countries have been targeting the wider definition of the money supply, correlating in line with their corresponding EMS partners.[4] In general, the wider the definition, the more direct the relationship between monetary growth and other economic variables, as it is of considerable importance in influencing the spending of individuals. For reasons of paucity of the available data, however, we choose the basic money stock M0/M1 instead of a broad money target. The data for announced monetary target M3 in Germany and France are only available from 1988Q1 (see IMF IFS, September 1991) and 1987Q1 (IMF IFS, July 1990) respectively, although the use of broad money is to be preferred, when available.

Figs. 6.3 and 6.4 respectively show the separate and joint time plots of LM0/LM1 for the Big Four over the period 1969Q2-1996Q4. Whilst LM0/LM1 rises approximately in line in all four of these series, there is a dramatic increase in Germany's LM1 in the first quarter of 1991. Rapid domestic private sector credit growth, primarily due to privatisations and government subsidies for construction and investment in eastern Germany, explains much of this growth. The German authorities did not revise their monetary target when they raised the discount rate, and they are currently suffering a significant target overshoot for all German money. There may, however, be good temporary reasons for this. The circulation of the D-Mark has risen in the east as the privatisation process proceeds.

Consumer price indices

All series used are seasonally adjusted, except for aggregate consumer prices and treasury-bill interest rates. Figs. 6.5 and 6.6 present individual graphs of LCPI for the major four industrialised countries in Europe over the period 1960Q1-1996Q3, and a joint one,

[4]The EMS, designed primarily to keep the member countries' exchange rates within the limits set, can be best construed as German-led, so that other countries peg to the D-Mark while Germany pursues.

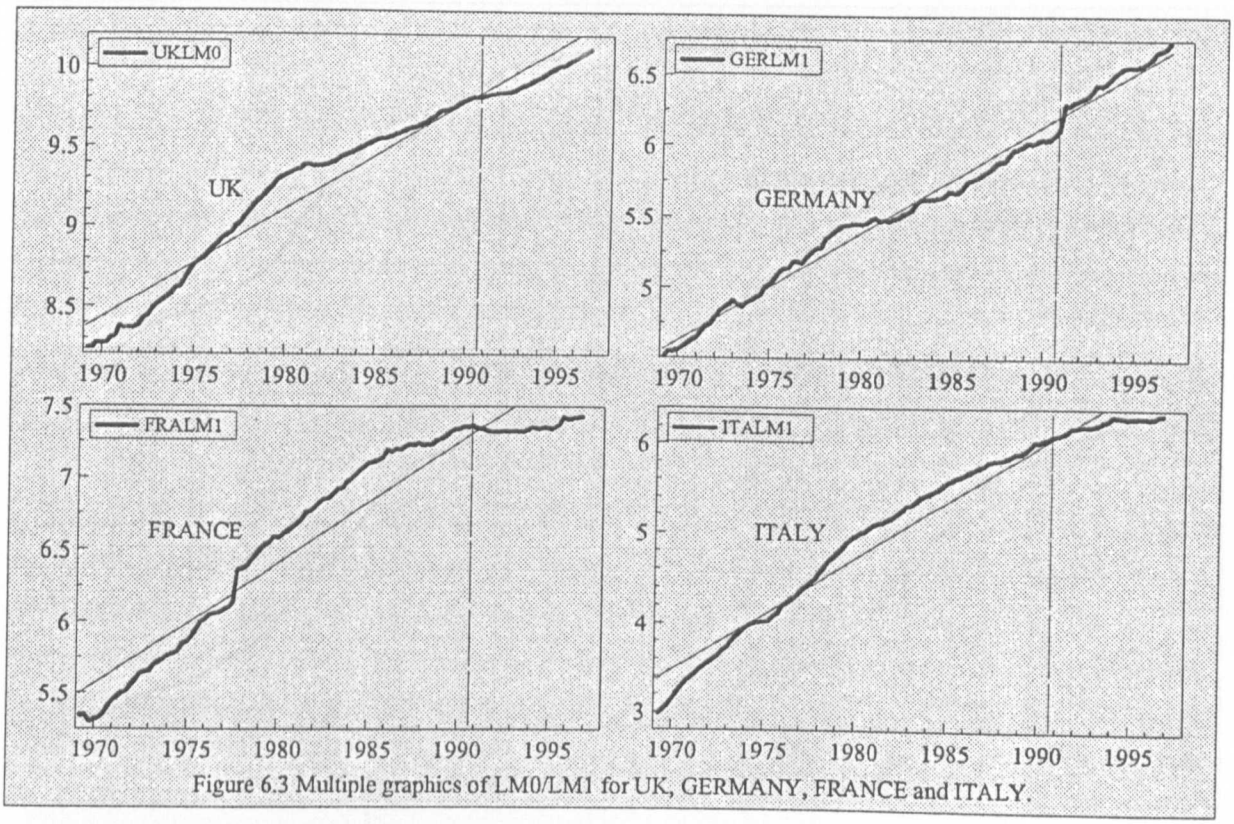

Figure 6.3 Multiple graphics of LM0/LM1 for UK, GERMANY, FRANCE and ITALY.

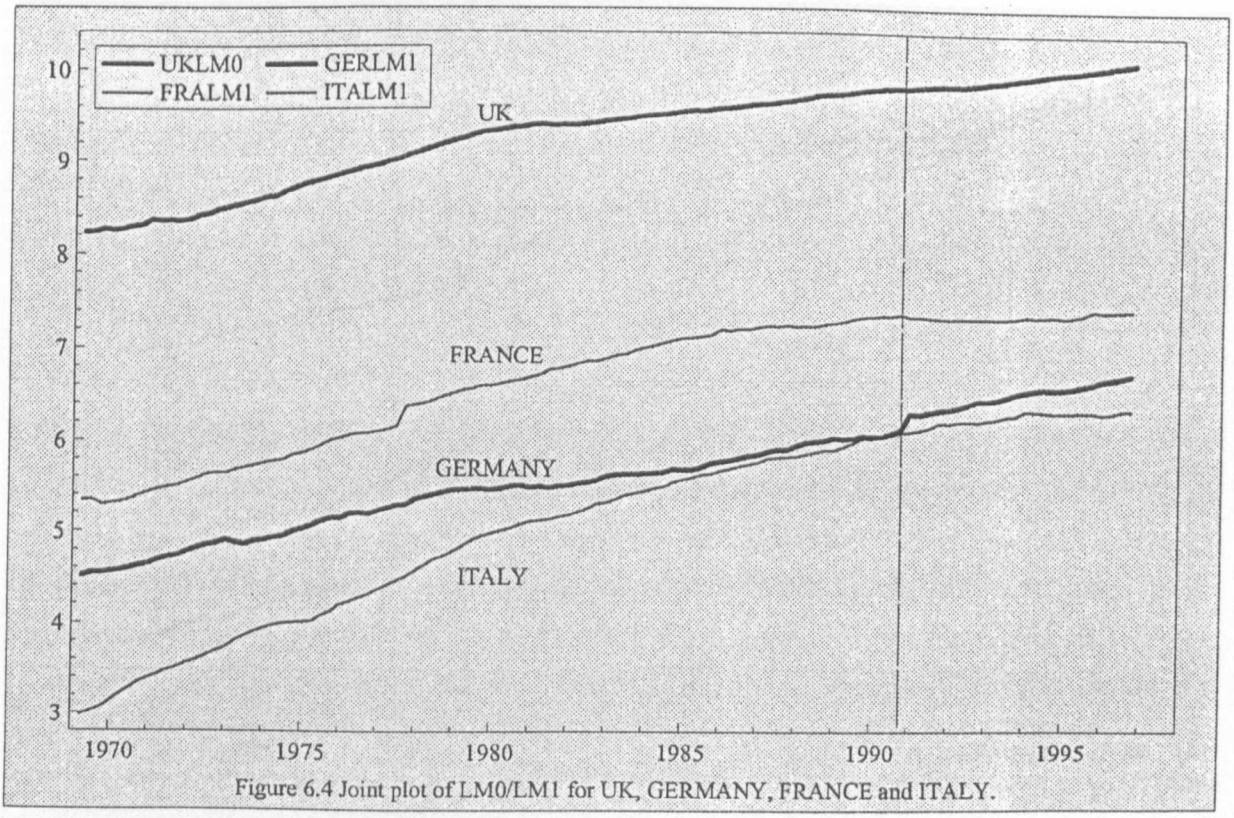

Figure 6.4 Joint plot of LM0/LM1 for UK, GERMANY, FRANCE and ITALY.

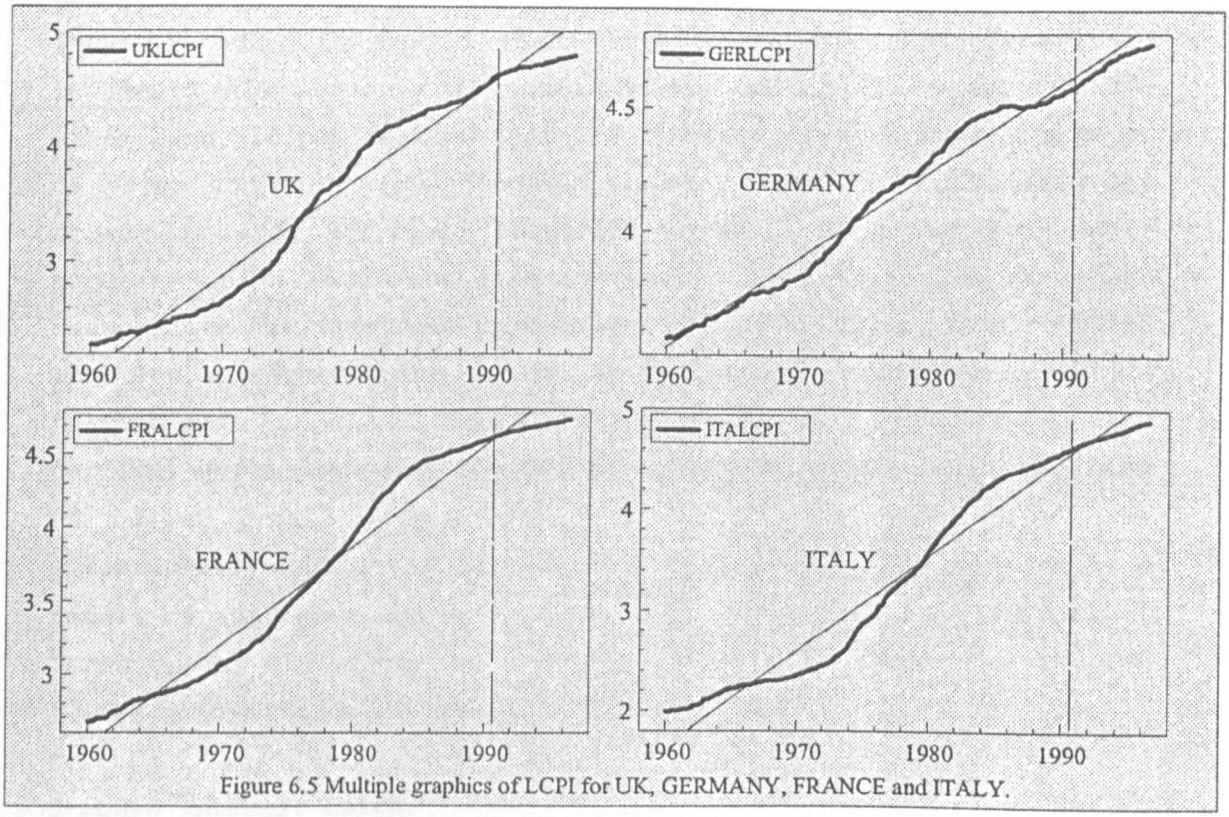

Figure 6.5 Multiple graphics of LCPI for UK, GERMANY, FRANCE and ITALY.

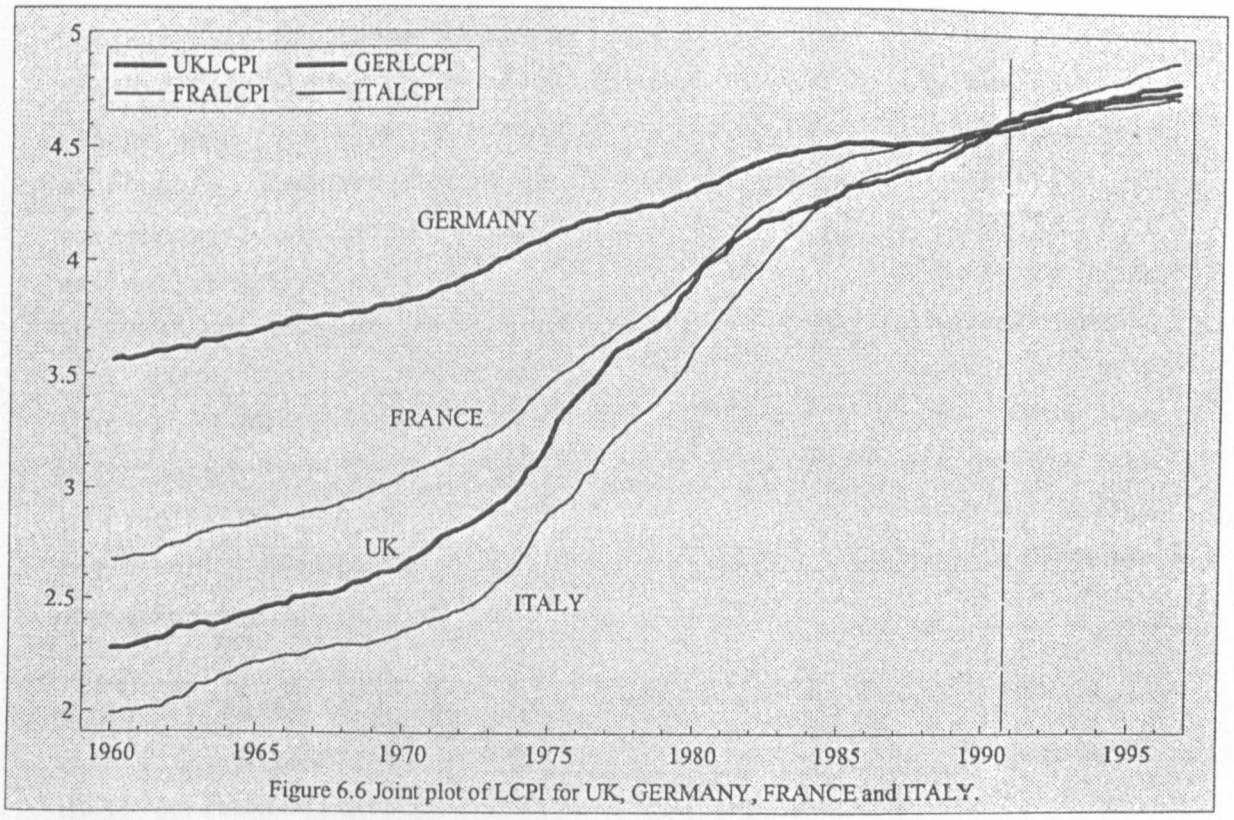

Figure 6.6 Joint plot of LCPI for UK, GERMANY, FRANCE and ITALY.

having a comparable vertical scale.

As far as individual price series are concerned, it can be seen that they all show an upward trend over time, but unadjusted price movements display only mild seasonality during the study. On the other hand, it can be seen that the cyclical fluctuations of a group of price series are synchronised across countries. Intuitively the reason is that if these countries were to form the core of a monetary union in Europe then eventually they would have to share essentially the same rate of consumer price inflation. This required that the aggregate price levels, as well as the aggregate inflation rates, had to adjust in different countries. Furthermore, the cyclical developments of the prices seem to be well matched by those of output growth within each country. In the early-1980s, the exceptionally strong GDP growth caused consumer prices to rise as a result of increased demand. In contrast, all prices fell in nominal terms during the trough of economic activity in 1990, contributing to the general deceleration in inflation rates.

3-month interest rates

Unlike the first three indicator series, the short interest rates of the Big Four appear to be almost trend-free in the long run. But, as can be seen from Figs. 6.7 and 6.8 over recent years, there has been a roughly downward trend in all four series in addition to coincidental movements between them. The underlying cause of this was policy commitments such as the decision to form a monetary union in 1999 and the need for interest rates in Europe to converge before that date. Interest differentials some 5 percent higher in Italy than in Germany indicate that a realignment of the Lira was anticipated. This policy, however, would not lead to such a big loss in credibility for any one country as would result from a unilateral adjustment. Hence, it would allow both the Italians to realign downwards and the Germans to realign upwards.

Moreover, as is shown in Fig. 6.9, over the past decade interest-rate fluctuations seem to be followed by the cyclical movements of their corresponding LGDP/LGNP after one or one-and-a-half years,

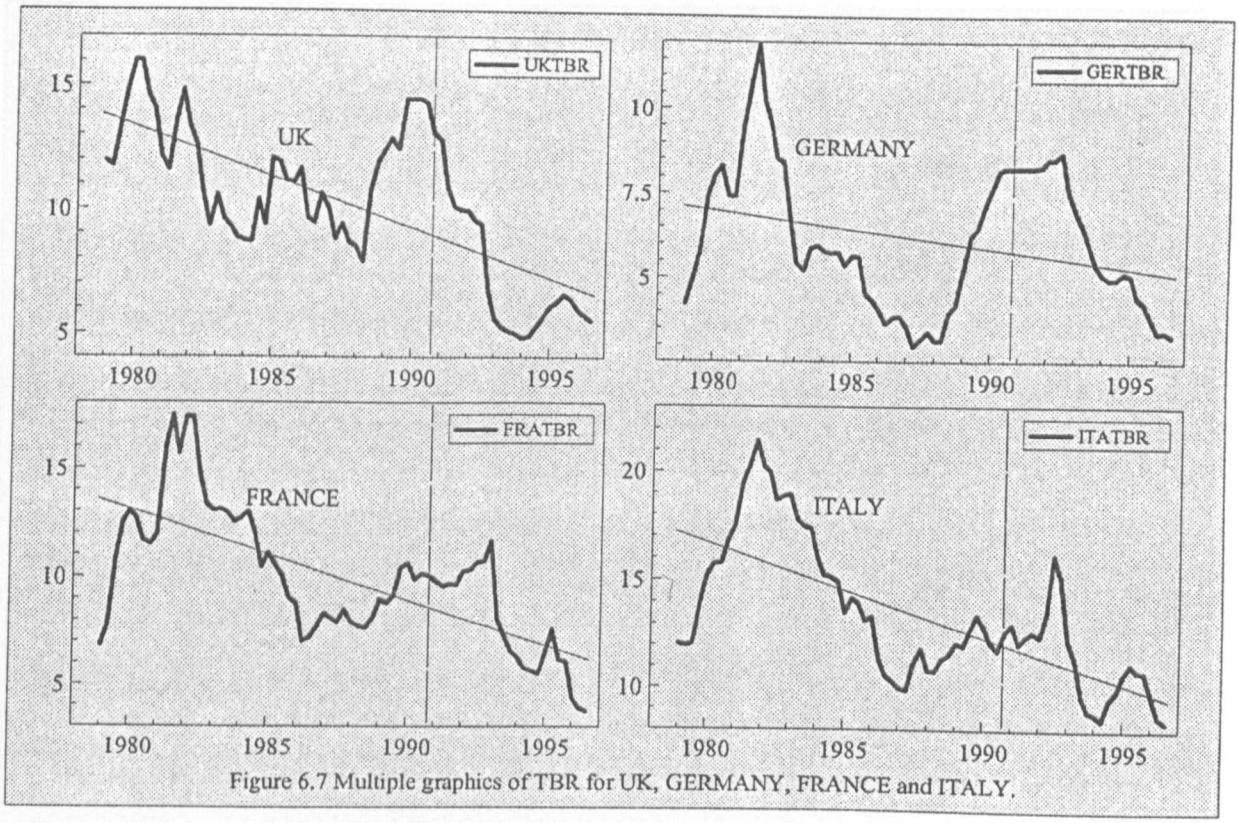

Figure 6.7 Multiple graphics of TBR for UK, GERMANY, FRANCE and ITALY.

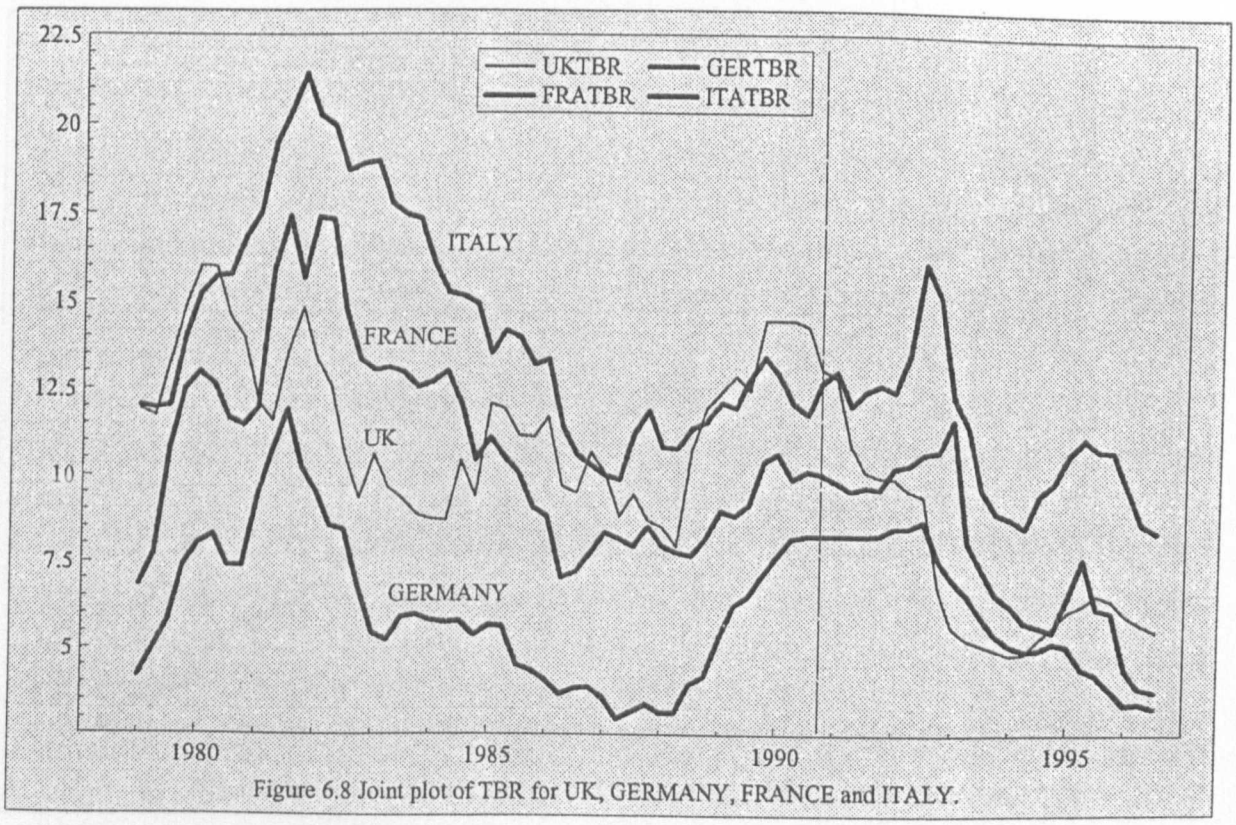

Figure 6.8 Joint plot of TBR for UK, GERMANY, FRANCE and ITALY.

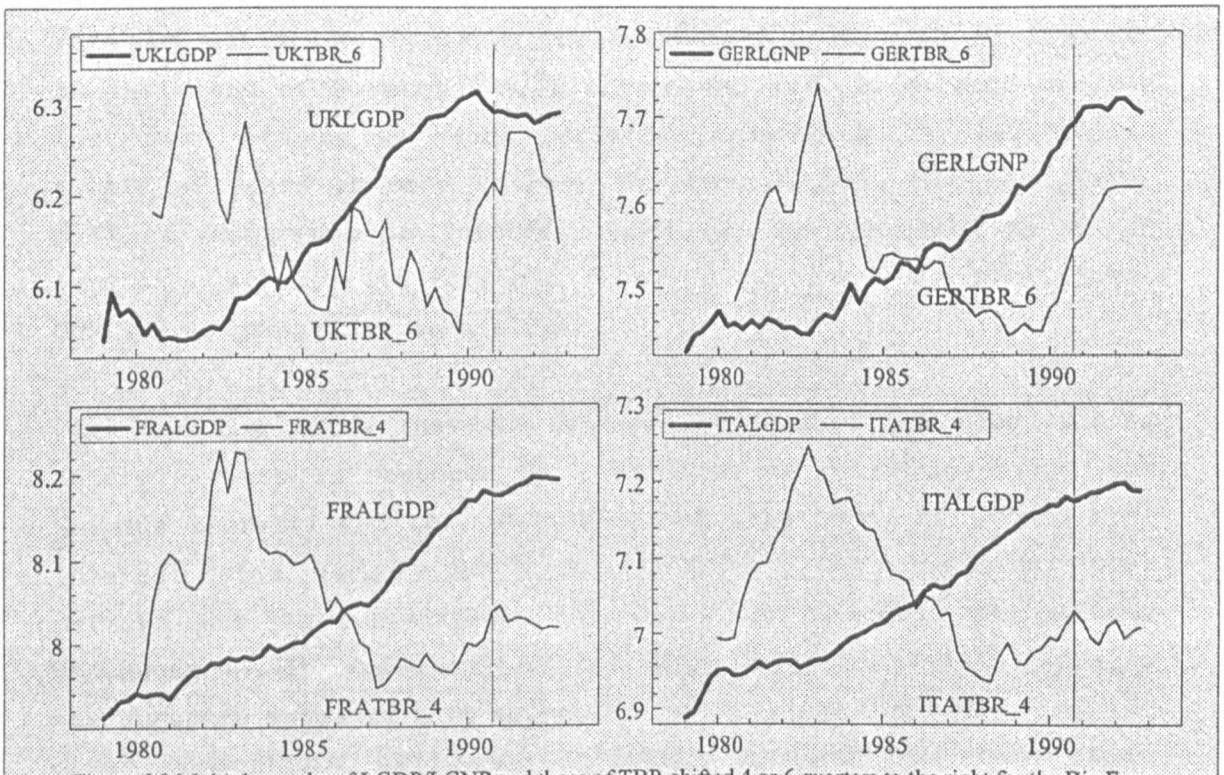

Figure 6.9 Multiple graphs of LGDP/LGNP and those of TBR shifted 4 or 6 quarters to the right for the Big Four.

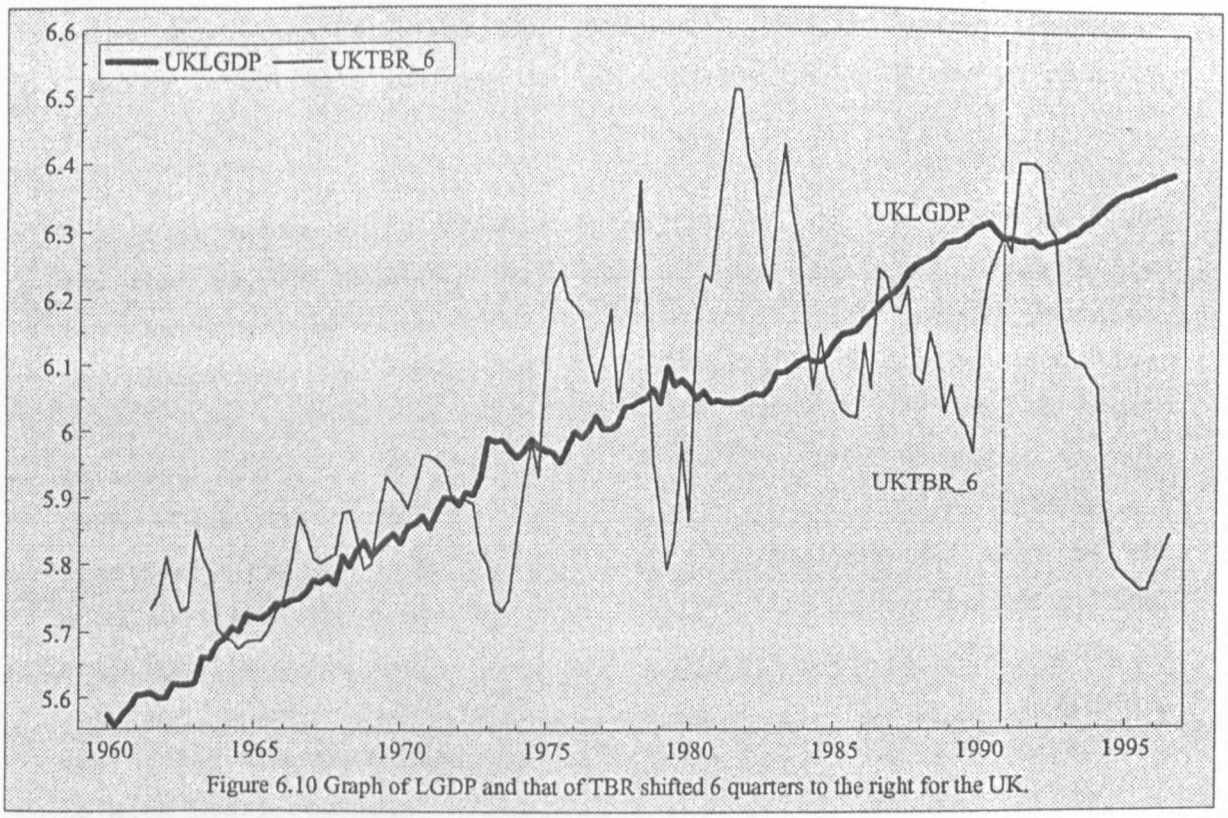

Figure 6.10 Graph of LGDP and that of TBR shifted 6 quarters to the right for the UK.

reflecting in part the policy stance of different governments by changes in interest rates in response to economic growth at different periods. Usually, the introduction of a high real interest rate may either be to dampen the overheated economy or be regarded as the signs of a recession, although the precise effects of this instrument are by no means perfectly understood. It may take a couple of years or more to have its full effects [George (1997), p.7]. A sharp cut in interest rates can be conceived of as an important stimulus to economic recovery and we would thus expect GDP/GNP to grow after such a cut.

Fig. 6.9 also shows that the British authorities had been cutting interest rates in response to a slowdown in activity, whilst the Germans probably had not. The sustained high level of German real interest rates, largely due to the pressure of the burgeoning fiscal deficit and its implications for the stock of debt, had serious consequences. The commitment of Sterling, the Franc and the Lira to the European exchange rate mechanism (ERM) of the EMS meant that the other competitors had also to keep their interest rates high vis-à-vis Germany. This did not suit the cyclical position of those countries, and the constraint on output growth of high real interest rates was exacerbated by the consequent loss of competitiveness. Clearly, sterling's leaving the ERM in September 1992 was germane to all of this.

Fig. 6.10 shows (to a roughly comparable scale) the graph of LGDP and that of TBR shifted 6 quarters to the right (i.e., the latter led the former by 6 quarters) for the UK over the last 30 years or so. Comparing the UK's LGDP and TBR, the historical negative correlation between the two stands out. That is, if the UK short interest rates are moved 6 quarters to the right relative to her LGDP along the time-axis, the two series will move inversely. Sometimes, however, the nature of what we observe may well be as much political as economic, as it may be largely generated by the artificially stimulated recovery before elections. The sharp cuts in interest rates in the hope they would aid the recovery indicate again that the government in office was willing to take any action to kick-start recovery. It also smacked of political desperation.

§6.3 Multi-Country Modelling(MCM) —— A Flexible Method of Hypothesis Testing on Popular Econometric Software Packages

The aim of this section is to proceed further along the six-step process with a focus on formulation of empirical BVAR and BVAR-EC forecasting models for the four major European economies. Several frequently requested tests associated with them will be considered successively. These are: (a) the treatment of the order of integration of each series involved; (b) the choice of the appropriate lag length of a finite-order VAR in levels; (c) the determination of stationary cointegrated linear combinations; and (d) the exploration of the best setting of the prior hyperparameters. In this study, all the necessary computations will be carried out using the available Microfit, MicroTSP and RATS software packages on a PC, with graphic and tabular displays for the final results.

6.3.1 Determination of Long-Run Equilibrium Relationships

In order to characterise the long-run properties of potentially nonstationary macroeconomic time series from the given set of data samples, we begin with an analysis of the order of integration of each series within the sample via the use of a one-sided ADF test statistic.

6.3.1a *Testing for Unit Roots or the Degree of Integration of Individual Series*

It is customary to test for the number of unit roots (or the orders of integration) of all the variables being analysed prior to testing for cointegration. Obviously, without stochastic trends there will be no common stochastic trends (CSTs), and without common trends there will be no cointegration. In this work, a series of ADF regressions (constructed along the lines of Dickey & Fuller 1979, 1981) is applied to each series separately to investigate whether or not the data are consistent with the hypothesis of one unit root at

the 5 percent level of significance.[5] To save space, only the results of the scalar ADF tests with different specified lags for the variables examined are reported in Table 6.1, rather than all the estimated equations.

Table 6.1: ADF unit root tests of individual series*

UK		LGDP	LMO	LCPI	TBR
Null hypothesis	I(1)	-0.79 (with trend)	-0.55 (with trend)	-1.45	-2.05
	I(2)	-5.13 (with trend)	-2.57 (with trend)	-2.63	-5.18
GERMANY		**LGNP**	**LM1**	**LCPI**	**TBR**
Null hypothesis	I(1)	-1.13	-0.82	-1.21 (with trend)	-1.96
	I(2)	-4.37	-5.03	-2.25 (with trend)	-4.06
FRANCE		**LGDP**	**LM1**	**LCPI**	**TBR**
Null hypothesis	I(1)	-1.43	-1.12	-1.17 (with trend)	-1.32
	I(2)	-3.35	-3.96	-2.73 (with trend)	-3.79
ITALY		**LGDP**	**LM1**	**LCPI**	**TBR**
Null hypothesis	I(1)	-0.78 (with trend)	-0.93 (with trend)	-0.53	-1.52
	I(2)	-5.19 (with trend)	-3.52 (with trend)	-3.51	-3.15

*Note: the 5 percent critical value of the ADF test (with drift) is -2.89 for T=100 (or -2.86 for T=∞) (Mackinnon 1991). When a (deterministic) time trend is included with a coefficient that appears to be significantly nonzero (under the null), the relevant ADF critical value will be -1.64 asymptotically, i.e. the usual lower-bound critical value of the standard normal N(0,1) distribution (Nankervis & Savin 1987, the bottom panel of Table 1). Otherwise, the associated critical value will be -3.45 instead for the sample sizes considered here (Mackinnon 1991).

[5] As Sims (1987, p.443) points out, there is no scientific justification for testing hypotheses at the 5% significance level in every application. Publishing such statistics and focusing on 5% levels is common because testing at fixed levels facilitates communication.

Here, in most cases, the above 't-ratio' unit root tests on the lagged levels are based upon the ADF regressions with a drift and with 4 lagged first differences over the sample through 1990Q4. The only exceptions are those tests of TBRs for Germany, France and Italy, and of the British LM0, Italian LM1 and LCPIs in all four countries. The former have a relatively short lag structure of 2 quarters due to the sample size restrictions, while the latter have time, denoted T, included in addition to an intercept, denoted INPT, as a regressor. For each level series, using the relevant 5 percent critical values, the estimates (given by the t-statistics) in Table 6.1 fail to reject the hypothesis that there is a unit root in the autoregressive representation. LCPIs, the UK's LM0 and Italy's LM1 will be tentatively treated as I(1) around a trend.[6] Thus, one plausible characterization of the long-run properties of the data consistent with the empirical evidence can be summarized in Table 6.2.

Table 6.2: One possible characterization of the trend properties of the data

	LGDP/LGNP	LM0/LM1	LCPI	TBR
UK	I(1)	I(1) (with trend)	I(1) (with trend)	I(1)
GERMANY	I(1)	I(1)	I(1) (with trend)	I(1)
FRANCE	I(1)	I(1)	I(1) (with trend)	I(1)
ITALY	I(1)	I(1) (with trend)	I(1) (with trend)	I(1)

It is important to reiterate that, although the ADF t-tests are widely used, care should be exercised in their use as the power of

[6] However, it is argued that the data need not be detrended. In a VAR, a trending variable will be well approximated by a unit root plus drift, and the drift will dominate the long-run behaviour of the nonstationary process [see Enders (1995), pp.300-1]. Moreover, a study by Hansen (1989) has shown that if the genuine data contain both deterministic and stochastic trends, removal of time trends from the data may worsen the performance of the normal asymptotics considerably.

these tests may be small, especially against trend-stationary alternatives. Three applied issues concerning the computation of the classical ADF tests need to be mentioned here.

Firstly, since the nonstandard (large and small sample) distributions of the test statistics are tabulated by simulation on unadjusted series, it is likely that test statistics computed using seasonally adjusted data will be severely biased in favour of the null of a unit root. Hence, if quarterly data are to be used, they should if possible not be seasonally adjusted. On this point see, *inter alia*, Davidson and Mackinnon (1993), p.714, but see all of Chapter 20. Unfortunately, seasonally unadjusted data for many important time series, like real GDP/GNP, are not available in the economies investigated here. In addition, the quarterly nature of seasonally unadjusted data may make it necessary to include a set of seasonal dummies and to account for as far back as fourth-quarter or eighth-quarter serial correlation. Even twelve quarters are not uncommon. Hence, the inclusion of data that exhibit strong seasonality into the information set will come at the potential cost of requiring many degrees of freedom and may lead to the problem of overfitting.

Secondly, it may well be that the ADF outcome is quite sensitive to the choice of lag length even for the same data (see Davidson & Mackinnon, op. cit., pp.710-5). One way of dealing with this would be to try a group of tests associated with various possible autocorrelations to ensure that the OLS residuals approximate an empirical white-noise process.

Finally, Christiano and Eichenbaum (1989) have demonstrated that, when deterministic nonstationarities are present, it is difficult or impossible to distinguish between deterministic and stochastic nonstationarities. In particular, structural breaks and regime shifts are likely to bias such tests toward the acceptance of a unit root [see Godfrey (ed) (1992), Chapter 1, especially pp.44-57]. For example, the failure to reject the unit root null in French LM1 may be the result of a Type II error. The series may be better expressed as trend-stationary, possibly about a one-time structural break at

the fourth-quarter of 1977.

Taking these factors into account, it may be said that testing for unit roots, while extremely important, is not without difficulties. The test results given in Table 6.2 are just one possibility. This characterization of the long-run properties of the data will also need to pass the test of time; that is, it needs to continue to hold outside the period of estimation.

6.3.1b *Testing for Cointegration Relationships for Those Variables That Are Integrated in Both Single- and Multi-Country Cases*

Once the univariate unit root tests have been completed, we move on to the estimation and testing of long-run cointegration relationships among subsets of the integrated (I(1)) variables, using Johansen's full-system maximum likelihood (ML) estimation methodology. However, prior to the determination of the cointegration rank, r, i.e. the number of independent CVs, the 'pre-test' specification of maximum lag truncation parameter, p, in the underlying VAR must also be made.

Since it is not feasible to model all the series concerned with a reasonably distributed lag jointly and the series are chosen for their economic importance, not for their statistical properties [Johansen (1996), p.34], in the following subsections we will mainly consider two types of quadrivariate systems: closed single-country and open cross-country ones. The former are various country models of the four different series in domestic economies (which are set to examine activities in individual countries), whilst the latter are systems of the four similar series across countries (which are used to characterize the dynamic interactions among the economies).[7]

[7] At the Johansen's estimation stage, the maximum number of endogenous I(1) variables associated with the student version of Microfit 3.0 (that is available at the time of research) is only 10 [see Pesaran, M.H. & B. Pesaran (1991), pp.175-7]. With advances in econometric literature and software, the limitation on that number has been released to 12 after the first draft of this thesis [Pesaran, M.H. & B. Pesaran (1997), pp.449-50].

(1) *Test for the value of the truncation lag p*

Commonly, a listing of the likelihood ratio (LR) tests associated with successively reducing the lag length of the VAR representation by one is an integral and crucial step in empirical multivariate cointegration studies.[8] To test systematically the null of a restricted VAR(l) against the alternative of a common unrestricted VAR(m), l<m, we need to run two VARs over the same sample period and save the residuals for their covariance matrices. This sample statistic has the asymptotic chi-square distribution with degrees of freedom equal to the number of restrictions in the system.[9]

Given degrees of freedom considerations, the reported work in this study was to take the same feasible upper bound of m=6 for each integer valued VAR lag order and thus set aside the first six entries as presample values. A sequential testing scheme in the VAR context was based on LR tests, which were adjusted along the lines of that advanced by Sims (1980a), with small-sample corrections, and carried out on RATS software package. The corresponding outcome of the log-likelihood of each modified test statistic in such a sequence for both single- and multi-country models are displayed in Tables 6.3 and 6.4, respectively.

---

[8]Practice has shown that it is necessary to work with an approximated 'truncated' version of what would otherwise generally be an infinite-order VAR [cf. Gali (1992), p.719]. Also, a low-order VAR system can, from a practical viewpoint, provide a reasonable approximation [cf. Söderlind & Vredin (1996), p.371-2].

[9]As Sims *et al.* (1990) assert, tests of the joint significance of additional lags in a VAR do not suffer from the existence of unit roots, even if the VAR is estimated in levels. Hence, the LR test statistic, along the lines of Sims' (1980a), can be applied to both stationary and nonstationary processes. See also Lütkepohl (1991a), pp.382-4.

Table 6.3: The likelihood ratio (LR) tests for the lag lengths of single-country models			
model	**Lag 1 vs lag m**	**LR test statistic**	**Signif**
UK	Lag 5 vs lag 6	Chi-Squared(16)=23.91	0.091
	Lag 4 vs lag 6	Chi-Squared(32)=63.00	0.001
	Lag 3 vs lag 6	Chi-Squared(48)=76.34	0.006
	Lag 2 vs lag 6	Chi-Squared(64)=92.12	0.012
GERMANY	Lag 5 vs lag 6	Chi-Squared(16)=12.06	0.740
	Lag 4 vs lag 6	Chi-Squared(32)=36.47	0.269
	Lag 3 vs lag 6	Chi-Squared(48)=54.27	0.248
	Lag 2 vs lag 6	Chi-Squared(64)=71.01	0.256
FRANCE	Lag 5 vs lag 6	Chi-Squared(16)=20.00	0.220
	Lag 4 vs lag 6	Chi-Squared(32)=33.83	0.379
	Lag 3 vs lag 6	Chi-Squared(48)=45.09	0.593
	Lag 2 vs lag 6	Chi-Squared(64)=56.17	0.746
ITALY	Lag 5 vs lag 6	Chi-Squared(16)=18.80	0.279
	Lag 4 vs lag 6	Chi-Squared(32)=38.51	0.199
	Lag 3 vs lag 6	Chi-Squared(48)=56.45	0.189
	Lag 2 vs lag 6	Chi-Squared(64)=66.89	0.378

Table 6.4: The likelihood ratio (LR) tests for the lag lengths of multi-country models

model	Lag 1 vs lag m	LR test statistic	Signif
LGDP	Lag 5 vs lag 6	Chi-Squared(16)=10.77	0.823
	Lag 4 vs lag 6	Chi-Squared(32)=32.93	0.421
	Lag 3 vs lag 6	Chi-Squared(48)=47.00	0.514
	Lag 2 vs lag 6	Chi-Squared(64)=65.18	0.435
LMON	Lag 5 vs lag 6	Chi-Squared(16)=13.77	0.616
	Lag 4 vs lag 6	Chi-Squared(32)=31.74	0.480
	Lag 3 vs lag 6	Chi-Squared(48)=49.33	0.420
	Lag 2 vs lag 6	Chi-Squared(64)=59.08	0.651
LCPI	Lag 5 vs lag 6	Chi-Squared(16)=18.63	0.288
	Lag 4 vs lag 6	Chi-Squared(32)=54.00	0.009
	Lag 3 vs lag 6	Chi-Squared(48)=78.52	0.004
	Lag 2 vs lag 6	Chi-Squared(64)=127.67	0.000
TBR	Lag 5 vs lag 6	Chi-Squared(16)=17.91	0.329
	Lag 4 vs lag 6	Chi-Squared(32)=27.86	0.676
	Lag 3 vs lag 6	Chi-Squared(48)=36.54	0.887
	Lag 2 vs lag 6	Chi-Squared(64)=46.92	0.946

The LR statistic can be viewed as a kind of residual analysis where residuals computed under the null are tested against those of a general alternative to see whether they include extra useful information which should not be omitted. Now, suppose we want to determine whether 4 lags are appropriate for all equations. If the restriction of a reduced number of lags is not binding, we would expect $\ln|\Sigma_u(4)|$ to be equal to $\ln|\Sigma_u(6)|$. Large calculated values of the statistic would mean the restriction of only 4 lags, or the imposition of a block of zero restrictions on lags 5 to 6, is binding. Hence, we can reject the null hypothesis that lag length =4. If the resulting LR test is less than χ^2 at a prespecified

significance level, we would not be able to reject the null of 4 lags being tested, i.e. the restriction is not binding. For a readable account of this property, see Enders (1995), pp.312-5.

According to the tables, in almost all applications encountered here, the minimum significance level of LR test statistic is around p=3. Most succeeding tests appear to be insignificant, however. Part of the problem results from the fact that the LR test is based on asymptotic theory which may not be very useful in small estimation samples. In general, we may adopt the parsimony principle on the selection of lag length, as longer lags may give rise to overfitting and thus result in poor out-of-sample forecasts. But unless we face severe data constraints, a lag order of at least 2 is necessary. Apart from this, specification of the Bayesian prior in certain ways will allow a more generous lag length than in an unrestricted VAR model. The selected maximum lags for both single- and multi-country systems will in most cases be a 1-year-worth (i.e. the periodicity of the data), as summarized in Table 6.5.

Table 6.5: The 'truncated' lag lengths of different VARs in levels

Name of the model		Truncation lag $\hat{p}$(LR)
Single-country models	UK	6
	GERMANY	4
	FRANCE	4
	ITALY	4
Multi-country models	LGDP	2
	LMON	4
	LCPI	4
	TBR	2

After a tentative moderate value of p has been specified, the determination of the number of cointegrating relations, r, can be obtained. Also, a VAR of order 2, 4 or 6 in levels implies a VECM representation of order 1, 3 or 5 if the series are cointegrated. All subsequent empirical inferences will depend on these selected lag lengths, and all relevant computer programs (saved in ordinary ASCII text files) can be referred to the Appendix D of this thesis.

In any chosen model, the choice of the length of consecutive lags will be an iterative process as well as other specification issues, and the primary selection criterion is forecasting accuracy. Hence, in a modelling exercise, less reliance should be placed on those indices of model adequacy that are used as 'guidelines' to model construction, and more emphasis should be given to the performance of models outside sample periods and against rival models. For extensive discussion of these issues see, in particular, Pesaran (1987, p.18) and Spencer (1993, pp.407-21).

(2) *Test for the rank of the cointegration space r*

Conditional on a given value of p, the number of CVs, r, can be determined sequentially by means of the Johansen (1989) maximal eigenvalue (λ_{max}) statistics. It has been found that Johansen (1988, 1991) trace (λ_{tr}) tests of the cointegration rank give very similar, but slightly higher, values of r. A variety of test results of the cointegration rank, together with their associated 5 percent critical values, for both specific- and cross-country models are presented in Tables 6.6 to 6.13.

Table 6.6: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the UK model*

Null	Alternative	Test Statistic	95% Critical Value
r = 0	r = 1	27.86	27.14
r ≤ 1	r = 2	19.47	21.07

*Note: (i) the estimation period is 1970Q4-1990Q4;
(ii) the maximum lag in VAR is 6. Estimation for trended variables, no trend in DGP.

Table 6.7: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the GERMANY model*

Null	Alternative	Test Statistic	95% Critical Value
r = 0	r = 1	34.76	27.14
r ≤ 1	r = 2	29.95	21.07
r ≤ 2	r = 3	4.34	14.90

*Note: (i) the estimation period is 1976Q3-1990Q4;
(ii) the maximum lag in VAR is 4. Estimation for trended variables, no trend in DGP.

Table 6.8: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the FRANCE model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	40.58	27.14
r ≤ 1	r = 2	18.22	21.07

*Note: (i) the estimation period is 1980Q1-1990Q4;
(ii) the maximum lag in VAR is 4. Estimation for trended variables, no trend in DGP.

Table 6.9: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the ITALY model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	26.99	27.14

*Note: (i) the estimation period is 1978Q2-1990Q4;
(ii) the maximum lag in VAR is 4. Estimation for trended variables, no trend in DGP.

Table 6.10: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the LGDP model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	28.94	27.14
r ≤ 1	r = 2	11.83	21.07

*Note: (i) the estimation period is 1971Q3-1990Q4;
(ii) the maximum lag in VAR is 2. Estimation for trended variables, no trend in DGP.

Table 6.11: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the LMON model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	44.59	27.14
r ≤ 1	r = 2	18.18	21.07

*Note: (i) the estimation period is 1970Q2-1990Q4;
(ii) the maximum lag in VAR is 4. Estimation for trended variables, no trend in DGP.

Table 6.12: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the LCPI model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	14.68	27.07

*Note: (i) the estimation period is 1961Q1-1990Q4;
(ii) the maximum lag in VAR is 4. Estimation for trended variables, with trend in DGP.

Table 6.13: Cointegration LR test based on Johansen maximal eigenvalue (λ_{max}) for the TBR model*			
Null	**Alternative**	**Test Statistic**	**95% Critical Value**
r = 0	r = 1	32.05	27.14
r ≤ 1	r = 2	15.50	21.07

*Note: (i) the estimation period is 1979Q3-1990Q4;
(ii) the maximum lag in VAR is 2. Estimation for trended variables, no trend in DGP.

It can be seen from Tables 6.6 to 6.13 that at the usual 5 percent significance level, data support the existence of just one cointegration relation in most situations. The only exceptions are those tests for Germany, Italy and the LCPI. A 5% level test for Germany shows the presence of two stationary cointegration combinations, while the latter two cases show no sample evidence of cointegration in either of the systems of four series.[10] There are, therefore, arguments for working within an integrated framework where LGDP/LGNP, LM0/LM1, LCPI and TBR are allowed to link explicitly together either in the main specific European economies or across the countries over the long run. The summary statistics of the estimated rank of the r-dimensional cointegration space for both the single- and multi-country models are reported in Table 6.14.

[10] More specifically, the λ_{max} statistics for Italy and the LCPI, 26.99 and 14.68, are below the respective 5 percent critical values, 27.14 and 27.07 and, thus, the hypothesis of noncointegration cannot be rejected.

Table 6.14: The estimated number of CVs for different models concerned		
Name of the model		Cointegration rank $\hat{r}(\lambda_{max})$
Single-country models	UK	1
	GERMANY	2
	FRANCE	1
	ITALY	0
Multi-country models	LGDP	1
	LMON	1
	LCPI	0
	TBR	1

Johansen's post estimation of cointegrating vectors $(\underline{\alpha})$, adjustment matrix $(\underline{\omega})$ and long-run matrix $(\underline{A}_p(1)=\underline{\omega}\underline{\alpha}')$, after an arbitrary normalization, can be found based on the relevant maintained cointegration rank.

(3) *Estimate cointegrating vectors and adjustment matrix, and save the residual vectors associated with the r CVs*

Given the dimension of the cointegration space, r, Johansen's normalised ML estimates of CVs and adjustment matrices of different four-equation models, except for Italy and the LCPI, are displayed in Tables 6.15 and 6.16.

Table 6.15: The ML estimates of normalised α and ω for the UK, GERMANY and FRANCE models				
UK	**LGDP**	**LMO**	**LCPI**	**TBR**
$\hat{\underline{\alpha}}$	1.000	0.295	-1.199	0.278
$\hat{\underline{\omega}}$	0.016	0.004	-0.014	1.160
GERMANY	**LGNP**	**LM1**	**LCPI**	**TBR**
$\hat{\underline{\alpha}}$	1.000	-0.286	0.086	-0.017
	1.000	**-1.590**	**2.871**	**-0.024**
$\hat{\underline{\omega}}$	0.024	-0.173	0.121	-9.451
	0.019	**0.033**	**0.026**	**1.979**
FRANCE	**LGDP**	**LM1**	**LCPI**	**TBR**
$\hat{\underline{\alpha}}$	1.000	-1.033	0.567	0.051
$\hat{\underline{\omega}}$	0.012	0.037	-0.034	1.936

Table 6.16: The ML estimates of normalised α and ω for the LGDP, LMON and TBR models

LGDP	UKLGDP	GERLGNP	FRALGDP	ITALGDP
$\hat{\underline{\alpha}}$	1.000	1.093	3.960	-5.360
$\hat{\underline{\omega}}$	-0.014	-0.053	-0.009	-0.053
LMON	**UKLM0**	**GERLM1**	**FRALM1**	**ITALM1**
$\hat{\underline{\alpha}}$	1.000	-2.872	-4.414	4.323
$\hat{\underline{\omega}}$	-0.009	0.004	-0.057	0.053
TBR	**UKTBR**	**GERTBR**	**FRATBR**	**ITATBR**
$\hat{\underline{\alpha}}$	1.000	107.804	-465.050	309.644
$\hat{\underline{\omega}}$	-0.000	-0.001	-0.002	0.000

It can be seen from Table 6.15 that there is a negative relationship holding between a cointegrated pair of series, LGDP and TBR, in both UK and France, but a positive one between the two model series, LGNP and TBR, in Germany. In addition, monetary aggregates go positively with consumer price indices through time in all three countries. Also, as demonstrated by Table 6.16, there is an explicit historical linkage amongst output growth, or money supply, series across the four countries. But for TBRs, the close long-run, or equilibrium, relationship appears only among Germany, France and Italy; UK's TBR is virtually excluded from the error-correction process. In particular, the ML estimate of the single CV $\underline{\alpha}$ for the former three is calculated to be (1.000, -4.243, 2.830)', with $\underline{\omega}$'=(-0.091, -0.213, 0.025). Hence, a TBR system of the four series should be equivalent to a partial VAR of the three with UK's TBR built on a single linear dynamic equation with no EC term.

After the Johansen's cointegration test results, the estimated residual vectors associated with the r CVs should be saved and tested for stationarity. Sometimes, the plotting of the residuals is an important diagnostic tool for checking stationarity. Only if all saved 'cointegrating residuals' are stationary, or I(0), will the lagged residuals, $\hat{\underline{\varepsilon}}_{t-p}\left(=\hat{\underline{\alpha}}'\underline{Y}_{t-p}\right)$, be fitted into each of the equations of a cointegrated VAR (VECM) as an additional regressor to

correct for short-run dynamics. Otherwise, the system should be built on a simple unrestricted VAR of order of p-1 in (first) differences.

6.3.2 Determination of the Best Scalar Prior Hyperparameter Settings

The present focus will be on revealing numerically the best scalar settings of the prior hyperparameters for all four-variable systems under scrutiny. According to the standard Minnesota prior, the choice of the best setting of the hyperparameters is consistent with the selection of the final version of a forecasting model, and both can be determined simultaneously. The final criterion for specifying hyperparameters over a certain range of values will be forecasting ability, rather than the usual specification tests or any other in-sample measures [see Spencer (1993), pp.409-11].

The sample period is divided into two non-overlapping sub-sample periods: a subperiod over which the model is estimated — the initial estimation period; and a subperiod over which forecasts are obtained and checked against actual data — the ex-post forecast period. If the anticipated future values of the series are reasonably well correlated with past episodes, we could expect that knowledge of the errors in the ex-post forecast period will allow us to modify our prediction of the error in the ex-ante forecast period. The ex-post, or simulated 'out-of-sample', forecasting performance will then be treated as an indicator for tuning the best choice of hyperparameters.

However, there exists a difficulty in using absolute measures of forecast accuracy such as the mean absolute percentage error (MAPE) or root mean square error (RMSE) to deliver rankings across models or forecasts, since such measures are not invariant under nonsingular, scale-preserving linear transformations for which linear models are invariant [see Clements & Hendry (1995), p.130]. Also, the 'optimality' of forecasts is usually defined, using absolute linear least squares, with respect to a given information set. Economic forecasters may not agree about the relevant

information set, which in the widest sense is in any case unknown, ex ante and ex post, and unmanageably large [see Wallis (1989), pp.43-4]. In the absence of an absolute standard, various comparative summary statistics have been developed and used for evaluating the relative accuracy of alternative forecasts. Among these, the informal use of Theil U statistics in the RATS software package is often preferred.[11]

The design of the experiment is as follows. We start with a common combination of ν=0.15 and w=0.5 for the overall tightness parameter, ν, and the tightness parameter on the 'other' variables for the relative tightness function f(i,j), w, in the highly restricted 2-parameter symmetric prior by default. For switching the prior from SYMMETRIC to GENERAL, a simple specification search procedure used in this application is to find weights $(\nu_j, w_{j1}, w_{j2}, w_{j3})$ on each row of the matrix of prior variances, by altering one while holding the others fixed, that minimize the Theil U values for each equation of the system. That is:

(a) run a system of unrestricted univariate OLS models by choosing ν_j=3.0 and the others zero to get benchmark Theil U's;

(b) run a system of restricted univariate models with a standard value for TIGHTNESS, ν_j=0.15. If the Theil U's in an equation become worse, then loosen up on the own lags by setting the diagonal element to 1.5 or 2.0; and

(c) run a standard SYMMETRIC prior for fractions w_{j1}, w_{j2}, w_{j3} = 0.5, respectively. If the Theil U's for one weight in an equation become worse, then tighten up on the 'other' variables by reducing

[11] Theil U statistic can be viewed as the ratio of the root mean square error (RMSE) of the forecast in question to the RMSE of the naive 'no-change' prediction from a random walk. When the value of Theil U's is less than unity, the forecasts of the model are judged to be more accurate than the naive forecasts. As a relative accuracy measure, Theil U statistic is likely to be superior to the use of absolute forecast error measures, such as the minimum MAPEs or RMSEs. For an entertaining account of the relevant issues, see Armstrong & Collopy (1992), and Fildes (1992) and the associated commentary.

the off-diagonal element to 0.1 or 0.01.

This would complete the selection of the hyperparameters as well as of the final forecasting model. The modeller might also examine changing patterns in Theil U values computed with successive overlapping two-, four-, six- or eight-step-ahead forecasts in a similar way before selecting the best setting of the hyperparameters. Such a procedure has the advantage of making the hyperparameter searches more objective and, in principle, statistically replicable. Moreover, an improvement over the traditional two-dimensional grid search strategy [Spencer (1993), pp.413-20] is that, in our modelling process, the search over values for the hyperparameters is straightforward and unbounded.

In terms of Leamer's (1982, 1991) sensitivity analysis, a study on the effect of a change in the prior covariance matrix $\underline{V}_j$ for the parameter vector β_j, j=1,...,4, is also carried out to see how sensitively our forecasts depend on departures of $\underline{V}_j$ from its 'best' choice $\underline{V}_{j0}$. This is normally done by entertaining a range of Bayes estimates given by the range of different values of prior covariance matrices around $\underline{V}_{j0}$: $\lambda^{-2}\underline{V}_{j0} \leq \underline{V}_j \leq \lambda^2\underline{V}_{j0}$; $\lambda>1$, which could be considered plausible in the light of the scientific context.

The final selections of the hyperparameters associated with the lowest values of the Theil U statistics for the two competing forecasting procedures developed: BVAR (specified in levels) and BVAR-EC (in differences), over the entire ex-post forecast period, are reported in Tables 6.17 and 6.18.[12]

[12]Space prevents us from exhibiting all findings here. Thus only the results obtained for the United Kingdom are presented. Other Tables 6.19 through 6.30 for both specific- and cross-country cases can be found in appendix E.

Table 6.17: Determination of the best hyperparameter setting (ν_j, w_{jk}) for UK BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGDP	0.96900 (.06,0,0,0)	0.92544 (.05,.5,0,0)	0.95482 (.05,1,.5,0)	0.73713 (.05,1,1,.5)
	0.96783 (**.05**,0,0,0)	0.91220 (.05,.8,0,0)	0.92884 (.05,1,.8,0)	0.67329 (.05,1,1,.8)
	0.96812 (.04,0,0,0)	**0.90443** (.05,**1**,0,0)	**0.92006** (.05,1,**1**,0)	**0.65727*** (.05,1,1,**1**)
LMO	0.50158 (.5,0,0,0)	0.48872 (.4,.06,0,0)	0.37624 (.4,.05,.1,0)	0.35828 (.4,.05,.01,.5)
	0.50095 (**.4**,0,0,0)	**0.48870** (.4,**.05**,0,0)	0.37204 (.4,.05,.05,0)	**0.35821*** (.4,.05,.01,**.6**)
	0.50097 (.3,0,0,0)	0.48873 (.4,.04,0,0)	**0.36958** (.4,.05,**.01**,0)	0.35824 (.4,.05,.01,.7)
LCPI	**0.60353** (3,0,0,0)	0.61475 (3,.5,0,0)	0.59945 (3,1,.09,0)	0.53707 (3,1,.08,.06)
	0.61340 (2.5,0,0,0)	0.60883 (3,.8,0,0)	**0.59923** (3,1,**.08**,0)	**0.53387*** (3,1,.08,**.05**)
	0.63078 (2,0,0,0)	**0.60709** (3,**1**,0,0)	0.59929 (3,1,.07,0)	0.53412 (3,1,.08,.04)
TBR	1.38981 (.1,0,0,0)	1.16896 (.01,.5,0,0)	1.00747 (.01,1,1,0)	0.98875 (.01,1,.9,.5)
	1.22655 (.05,0,0,0)	1.16760 (.01,.8,0,0)	**1.00720** (.01,1,**.9**,0)	0.98217 (.01,1,.9,.8)
	1.10833 (**.01**,0,0,0)	**1.16644** (.01,**1**,0,0)	1.00727 (.01,1,.8,0)	**0.97920*** (.01,1,.9,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.18: Determination of the best hyperparameter setting (ν_j, w_{jk}) for UK BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGDP	1.07926 (.1,0,0,0)	1.07234 (.01,.5,0,0)	1.121010 (.01,1,.5,0)	0.752811 (.01,1,1,.1)
	1.05810 (.05,0,0,0)	1.07223 (.01,.8,0,0)	1.121002 (.01,1,.8,0)	0.752808 (.01,1,1,.05)
	1.05200 (**.01**,0,0,0)	**1.07213** (.01,**1**,0,0)	**1.120996** (.01,1,**1**,0)	**0.752807*** (.01,1,1,**.01**)
LMO	0.75969 (.5,0,0,0)	0.77780 (.4,.5,0,0)	0.93285 (.4,1,.5,0)	0.78767 (.4,1,1,.03)
	0.75350 (**.4**,0,0,0)	0.77248 (.4,.8,0,0)	0.88597 (.4,1,.8,0)	**0.78729*** (.4,1,1,**.02**)
	0.75724 (.3,0,0,0)	**0.76919** (.4,1,0,0)	**0.86293** (.4,1,**1**,0)	0.79397 (.4,1,1,.01)
LCPI	**0.44583** (3,0,0,0)	0.44783 (3,.5,0,0)	0.39865 (3,1,.04,0)	0.30693 (3,1,.03,.12)
	0.45096 (2,0,0,0)	0.44670 (3,.8,0,0)	**0.39680** (3,1,**.03**,0)	**0.30651*** (3,1,.03,**.11**)
	0.48625 (1,0,0,0)	**0.44640** (3,**1**,0,0)	0.39807 (3,1,.02,0)	0.30665 (3,1,.03,.10)
TBR	**0.48142** (3,0,0,0)	0.41241 (3,.1,0,0)	0.39236 (3,.01,.15,0)	0.39910 (3,.01,.14,.5)
	0.48214 (2,0,0,0)	0.40027 (3,.05,0,0)	**0.39233** (3,.01,**.14**,0)	0.39458 (3,.01,.14,.8)
	0.48787 (1,0,0,0)	**0.39958** (3,**.01**,0,0)	0.39239 (3,.01,.13,0)	**0.39370*** (3,.01,.14,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Because of the shortage of data on German real GNP (available only through 1992Q4), the missing values must themselves be predicted first in both GERMANY and LGDP models. The ex-post, or inside-sample, forecast period for the variable concerned will therefore be 1991Q1-1992Q4. The period commencing with 1993Q1 will be treated as ex-ante, or outside-sample, period. In this regard, the practical forecasts made on 'future' values of German GNP from these two models are conditional rather than unconditional and are not endowed

with the same information set as other formal, model-based forecasts.[13]

Once the hyperparameters that seem to lead to the best forecasting model have been chosen, they are usually re-evaluated only every few years. More frequent updating of the hyperparameters is expensive, and experience in searching over the best settings of the hyperparameters has shown that it would probably yield little gains in the average accuracy of the forecasts. Although the hyperparameters are infrequently revised, the coefficients associated with the chosen hyperparameters are routinely updated using the Kalman filter (KF) estimation method each time, as a new data point becomes available.[14]

The Minnesota system of prior information or beliefs is not simple to use, but it does give forecasters a flexible way to express personal beliefs and an objective procedure for combining those beliefs with historical data to produce forecasts. In that sense, it represents a move away from traditional forecasting procedures towards comparatively cheap Bayesian procedures that seemed, by the late 1970s, more promising to at least some economists [see Todd (1984), pp.28-9].

§6.4 Conclusion

So far, we have developed a systematic six-step BVAR and/or BVAR-EC modelling procedure that can be readily carried out using the popular econometric software packages. To implement such a

[13] For a thorough discussion on the distinction between ex-post and ex-ante forecasts, as well as that between conditional and unconditional forecasts, see Pindyck & Rubinfeld (1991), Chapter 8.

[14] Here, the KF algorithm is applied in this way: ESTIMATE initializes the KF over the pre-forecast period, then the loop executes each KALMAN command to update continually the estimates till the time of the forecast period as new observations become available. This opens the way for estimating and evaluating models where coefficients are allowed to vary over time and makes it highly useful for model specification and prediction [see Hall (1993), pp.95-7].

procedure, the available quarterly time-series data used in estimations and analyses are obtained from various currently existing IMF, OECD and ONS databanks through MIDAS. Recent advances relating to unit roots, lag structures and cointegrated relations are tested in sequence in both specific- and cross-country cases. The best settings of the prior hyperparameters associated with the final forecasting models are revealed, making informal use of Theil U statistics. In the next chapter, the resulting BVAR models that incorporate both prior and cointegration restrictions will be estimated and used for forecasting the main European economies, utilizing the Kalman filter recursions.

CHAPTER 7

FORMULATION, ESTIMATION AND EVALUATION OF MACRO BVAR FORECASTING MODELS FOR THE EUROPEAN ECONOMIES

§7.1 Introduction

The main objective of this chapter is to use the models constructed as forecasting devices to obtain quarterly forecasts of selected variables for the leading European economies and to assess the gains for forecast accuracy from imposing prior and long-run constraints. The empirical setting-up process of K=4-dimensional BVAR, and also BVAR-EC, forecasting models based on the chosen prior hyperparameter settings and the reported ML estimates of $\underline{A}_p(1)=\underline{\omega}\underline{\alpha}'$ consists of three stages. First, the r lagged stationary residual series, $\hat{\underline{\varepsilon}}_{t-p}\left(=\hat{\underline{\alpha}}'\underline{Y}_{t-p}\right)$, from the previous estimation are included unrestrictedly in each equation of the system as additional regressors if there is cointegration. Next, the multivariate normal prior distributions are imposed on the other parameters of the system (except for the constant and error-correction terms). Finally, the coefficients of the resulting four-equation dynamic system satisfying both prior and cointegration constraints are estimated jointly using the mixed estimation technique proposed by Theil (1971, pp.347-52). It is hoped that, through this work, we will be able to generate useful ex ante forecasts; to promote an understanding of different national economies investigated; and to capture interdependencies among European economies.

§7.2 Estimated Version of BVAR and BVAR-EC Forecasting Models in Matrix Notations — An Appropriate Procedure for Sequentially Updating the Models' Parameter Estimates

In retrospect, we know that a typical dynamic forth-order

four-dimensional BVAR forecasting model can be given by:

$$Y_{jt} = C_j + \sum_{n=1}^{4}\sum_{i=1}^{4} A_{jn,i} Y_{n,t-i} + u_{jt}; \qquad (7.1)$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \quad 0.01 \leq \nu_j \leq 3.0; \quad 0.01 \leq w_{jk} \leq 1.0;$$

$$j, k = 1, \ldots, 4; \quad t = 1, \ldots, T.$$

Here, (7.1) is a set of four recursive equations specified in levels, in which the unknown parameter values are to be replaced by estimates; and

ν_j and $\nu_j w_{jk}$ are the hyperparameters of the prior distributions imposed on the lagged coefficients of the jth-equation of the system.

Litterman (1986a) specifically advises against taking first differences even if the variables contain a unit root, since information concerning the comovements in the data is thrown away. He argues that the goal of a VAR analysis is to determine the important interrelationships among the variables, not the parameter estimates. This is especially good advice when constructing a Bayesian VAR forecasting model with unit roots, since the belief that the series has a unit root in its autoregressive representation is easily incorporated in the stochastic prior restrictions centred about a random walk (plus drift) process. [See, for example, Enders (1995), p.301, and Spencer (1993), p.411.]

In contrast, if cointegration exists, its third-order four-dimensional reduced form BVAR-EC model can be written as:

$$\nabla Y_{jt} = C_j + \sum_{n=1}^{4}\sum_{i=1}^{3} A_{jn,i}^{\dagger} \nabla Y_{n,t-i} - \sum_{s=1}^{r} \hat{\omega}_{js} \hat{\varepsilon}_{s,t-4} + u_{jt}; \qquad (7.2)$$

$$u_{jt} \sim NID(0, \sigma_{jj}^2); \quad 0.01 \leq \nu_j \leq 3.0; \quad 0.01 \leq w_{jk} \leq 1.0;$$

$$j, k = 1, \ldots, 4; \quad t = 1, \ldots, T;$$

$$\hat{\varepsilon}_{s,t-4} = \hat{\underline{\alpha}}_s' \underline{Y}_{t-4} = \sum_{i=1}^{4} \hat{\alpha}_{si} Y_{i,t-4}; \quad s = 1, \ldots, r; \qquad (7.3)$$

$$Y_{jt} = Y_{j,t-1} + \nabla Y_{jt}; \quad j = 1, \ldots, 4. \qquad (7.4)$$

Here, (7.2) is a group of four estimated recursive equations in

first differences;

(7.3) is the lagged estimated cointegrating residuals, denoted EC(-p), included in each of the four equations as an exogenous variable at full weight along with the intercept term. It is then computed and updated at each forecast step to influence continually the forecasts of ∇Y_{jt}, j=1,...,4, assuming that the underlying equilibrium relationships defined by the r estimated CVs $\hat{\underline{\alpha}}$ can be carried over into the forecast interval;

ν_j and $\nu_j w_{jk}$ are the prior hyperparameters specified and imposed on all the coefficients except the constant and error-correction terms; and

(7.4) is a collection of four identities required to transform the four forecasted difference series into the needed level series, which will then be used to calculate the EC term in subsequent forecasts.

Based upon these fundamental equations, a selection of the empirical results of estimated coefficients, t and R^2 statistics for the proposed two competing quarterly forecasting systems of equations 1 to 4 for the UK can be summarized as in Tables 7.1 and 7.2 below.[1]

[1]Here, only the results and analyses for the UK are provided, otherwise it would be excessively burdensome to report them all. Similar analyses apply as well to the other cases.

Table 7.1: Mixed estimates of the final BVAR forecasting model for the UK

Dependent variable	Eq.1 LGDP	Eq.2 LMO	Eq.3 LCPI	Eq.4 TBR
Constant	0.2797 (1.17)	0.0349 (0.47)	-0.5991a (-2.47)	-6.4123 (-0.73)
$LGDP_{-1}$	0.9744b (30.04)	0.0050 (0.36)	0.0535 (0.68)	0.3847 (0.41)
$LGDP_{-2}$	0.0024 (0.11)	0.0012 (0.16)	-0.2190a (-2.41)	0.0864 (0.18)
$LGDP_{-3}$	0.0016 (0.10)	0.0005 (0.11)	0.2211a (2.42)	0.0321 (0.10)
$LGDP_{-4}$	0.0013 (0.11)	0.0003 (0.07)	0.0187 (0.21)	0.0123 (0.05)
$LGDP_{-5}$	0.0014 (0.14)	0.0003 (0.09)	-0.1836a (-2.05)	0.0040 (0.02)
$LGDP_{-6}$	0.0008 (0.11)	0.0002 (0.06)	0.1098 (1.43)	0.0035 (0.02)
LMO_{-1}	-0.0074 (-0.18)	1.1551b (12.61)	0.2003b (3.16)	0.5032 (0.49)
LMO_{-2}	-0.0105 (-0.35)	0.0156 (0.13)	-0.0051 (-0.07)	0.0984 (0.17)
LMO_{-3}	-0.0047 (-0.22)	-0.1282 (-1.32)	-0.0310 (-0.57)	0.0228 (0.06)
LMO_{-4}	-0.0041 (-0.26)	-0.0223 (-0.28)	-0.0392 (-0.92)	0.0055 (0.02)
LMO_{-5}	-0.0014 (-0.11)	0.0589 (0.91)	-0.0160 (-0.45)	-0.0010 (-0.00)
LMO_{-6}	0.0004 (0.04)	-0.0859 (-1.73)	-0.0088 (-0.29)	-0.0002 (-0.00)
$LCPI_{-1}$	0.0023 (0.06)	-0.0002 (-0.04)	1.1550b (11.45)	-0.5028 (-0.55)
$LCPI_{-2}$	0.0077 (0.29)	-0.0000 (-0.01)	-0.2737a (-1.98)	-0.0962 (-0.17)
$LCPI_{-3}$	0.0050 (0.27)	-0.0000 (-0.00)	-0.1661 (-1.28)	-0.0340 (-0.09)
$LCPI_{-4}$	0.0037 (0.27)	-0.0000 (-0.02)	0.6154b (4.82)	-0.0081 (-0.03)
$LCPI_{-5}$	0.0036 (0.32)	-0.0000 (-0.03)	-0.6157b (-4.62)	0.0064 (0.03)
$LCPI_{-6}$	0.0037 (0.40)	-0.0000 (-0.03)	0.1973a (2.38)	0.0053 (0.03)
TBR_{-1}	-0.0005 (-1.24)	-0.0007 (-1.02)	0.0016b (2.91)	0.9957b (96.42)
TBR_{-2}	-0.0001 (-0.50)	0.0001 (0.08)	0.0002 (0.52)	-0.0012 (-0.23)
TBR_{-3}	-0.0001 (-0.55)	0.0002 (0.36)	0.0001 (0.21)	-0.0005 (-0.15)
TBR_{-4}	-0.0000 (-0.19)	0.0001 (0.23)	-0.0000 (-0.02)	-0.0003 (-0.11)
TBR_{-5}	-0.0000 (-0.27)	-0.0001 (-0.43)	0.0002 (0.84)	-0.0002 (-0.07)
TBR_{-6}	-0.0000 (-0.12)	-0.0001 (-0.33)	0.0001 (0.49)	-0.0001 (-0.05)
$\nu_j, w_{j1}, w_{j2}, w_{j3}$	.05,1,1,1	.4,.05,.01,.6	3,1,.08,.05	.01,1,.9,1
R^2	0.9880	0.9994	0.9998	0.7654
SEE	0.0135	0.0108	0.0084	1.3683
SSR	0.0146	0.0093	0.0057	149.7889
T	104			
Estimation period	1970:4-1996:3			

Note: figures given in parentheses denote t-ratios; a and b indicate statistically significant at 5% and 1% significance levels, respectively.

Table 7.2: Mixed estimates of the final BVAR-EC forecasting model for the UK

Dependent variable	Eq.1 ∇LGDP	Eq.2 ∇LMO	Eq.3 ∇LCPI	Eq.4 ∇TBR
Constant	0.0206 (1.21)	0.0151 (0.93)	-0.0132 (-0.71)	6.7896[a] (2.25)
EC_{-6}	-0.0021 (-0.90)	-0.0016 (-0.67)	0.0025 (0.93)	-1.0521[a] (-2.28)
$\nabla LGDP_{-1}$	-0.0014 (-0.13)	0.0258 (0.28)	0.0679 (0.79)	-0.5014 (-0.19)
$\nabla LGDP_{-2}$	-0.0000 (-0.00)	0.0399 (0.49)	-0.1365 (-1.55)	0.0644 (0.05)
$\nabla LGDP_{-3}$	-0.0000 (-0.01)	0.0345 (0.48)	0.1227 (1.34)	-0.0127 (-0.01)
$\nabla LGDP_{-4}$	-0.0000 (-0.02)	-0.0246 (-0.41)	0.1285 (1.44)	0.0169 (0.03)
$\nabla LGDP_{-5}$	-0.0000 (-0.00)	0.0209 (0.40)	-0.1198 (-1.38)	-0.0165 (-0.03)
∇LMO_{-1}	0.0010 (0.07)	0.2374[a] (2.41)	0.1348[a] (2.28)	4.6483 (0.39)
∇LMO_{-2}	-0.0001 (-0.02)	0.2729[b] (2.85)	0.0367 (1.01)	12.1911 (1.06)
∇LMO_{-3}	0.0000 (0.00)	0.0053 (0.06)	0.0119 (0.47)	3.7622 (0.36)
∇LMO_{-4}	-0.0001 (-0.04)	-0.0590 (-0.79)	0.0007 (0.04)	5.1853 (0.57)
∇LMO_{-5}	-0.0001 (-0.04)	0.1251 (1.92)	0.0006 (0.04)	0.4262 (0.05)
$\nabla LCPI_{-1}$	-0.0021 (-0.17)	-0.0418 (-0.41)	0.4612[b] (4.60)	-7.0481 (-0.47)
$\nabla LCPI_{-2}$	-0.0004 (-0.06)	-0.0258 (-0.29)	0.0807 (0.87)	13.7260 (1.08)
$\nabla LCPI_{-3}$	-0.0001 (-0.03)	0.1406 (1.80)	-0.1494 (-1.60)	1.5351 (0.12)
$\nabla LCPI_{-4}$	-0.0001 (-0.03)	0.0585 (0.85)	0.5391[b] (5.82)	-17.2312 (-1.34)
$\nabla LCPI_{-5}$	-0.0001 (-0.02)	-0.0104 (-0.17)	-0.2798[b] (-2.88)	20.6582 (1.66)
∇TBR_{-1}	-0.0000 (-0.00)	-0.0000 (-0.07)	0.0023[b] (2.96)	-0.0633 (-0.57)
∇TBR_{-2}	0.0000 (0.00)	-0.0000 (-0.03)	0.0012 (1.71)	-0.1151 (-0.95)
∇TBR_{-3}	-0.0000 (-0.00)	-0.0000 (-0.00)	0.0007 (1.27)	-0.1535 (-1.20)
∇TBR_{-4}	0.0000 (0.00)	0.0000 (0.02)	-0.0005 (-1.00)	-0.2138 (-1.59)
∇TBR_{-5}	-0.0000 (-0.00)	-0.0000 (-0.02)	0.0006 (1.41)	-0.1650 (-1.29)
$\nu_j, w_{j1}, w_{j2}, w_{j3}$	.01,1,1,.01	.4,1,1,.02	3,1,.03,.11	3,.01,.14,1
R^2	0.0118	0.3677	0.6803	0.2203
SEE	0.0142	0.0107	0.0093	1.2304
SSR	0.0160	0.0090	0.0068	119.5946
T	104			
Estimation period	1970:4-1996:3			

Note: figures given in parentheses denote t-ratios; **a** and **b** indicate statistically significant at 5% and 1% significance levels, respectively.

For comparison, the point estimates for both BVAR and BVAR-EC forecasting models were made (at the end of the maximum possible

sample) by mixed-estimation for the same set of variables over the same estimation sample period. It may be seen from Table 7.1 that, among the four linear estimated equations of the 'final' quarterly BVAR forecasting model, the estimates recovered from the whole sample for all but UK's LCPI appear to be a random walk process. The LCPI may be better expressed as a random walk about a trend process.

It can be seen from Table 7.2 that among the four estimated equations of the reduced form BVAR-EC forecasting model, both ∇LGDP and ∇TBR behave more or less like a 'white noise' process. The constant as well as the error-correction term appears to be statistically significant for UK's ∇TBR within the system. This confirms not only the standard prior for nonstationary series which has one unit root but also the standard white noise-like prior for stationary series. Since these variables are likely to be close to white noise, they may be easily buried in and hardly picked out from the disturbance terms. ∇LCPI and ∇LM0, however, show statistically significant dependence on the past. ∇LCPI also shows a significant relationship with the first lag of ∇LM0 and of ∇TBR.

Many of the parameter estimates of the models turn out to be statistically insignificant at conventional significance levels. The decision as to whether the 'true' parameter values are really zero is sometimes of crucial importance. The finding of a t-statistic within some given distance of zero might mean that the coefficient is really zero, or it could simply be that the data do not contain enough information to show convincingly that the coefficient is significantly different from zero. This problem is exacerbated in the case of existence of serious multicollinearity, because of the increased probability of type II error. Hence, the t-tests on individual regression coefficients may not be reliable guides for 'paring down' the model. [For an excellent and lucid account of the analysis, see Steward (1986), pp.129-34.]

A slightly different approach to the significance test may be adopted. Instead of using the critical value for a particular significance level as a definite boundary between acceptance and rejection of the null hypothesis, the reported statistics are viewed

as an indicator of the performance of a particular regressor in explaining the dependent variable, especially when the presence of multicollinearity is suspected. In taking such decisions, some weight should be given to any prior beliefs concerning the role of the variable in question and also to the likely costs of misspecification that would result from taking the wrong decision. If there is reason to believe that a variable was relevant to the explanation of the behaviour of the dependent variable, a small value of t-ratio might indicate the need to experiment further, rather than be as definite evidence that the corresponding coefficient is zero. This would be especially true of a variable that could be subject to very large changes. In such a situation, improperly treating even small nonzero values of coefficients as if they were zero could lead to rather serious errors when the model is used for forecasting or policy simulation.

As shown in Table 7.2, the estimated coefficients of error-correction terms, EC(-6), in all but ∇LCPI equations appear to be 'desired' negative. At the same time, the coefficient attached to the EC term in ∇LCPI equation appears to be 'unwanted' positive. The appearance of the positive sign in ∇LCPI could suggest that a given cointegrating relationship in the past may not be appropriate for use in this equation much beyond its range of estimation. It could also suggest that the specification of the model is appropriate, but that many coefficients have evolved through time due to structural breaks and regime shifts. If this is the case, it may be a caution against extending the so-called 'long-run' cointegrating relationships for real-time forecasts farther into the future.

In what follows, the experiment-based BVAR as well as BVAR-EC versions will be used in both individual- and multi-country contexts to produce a series of one-step-ahead quarterly forecasts for the principal western European economies. Certain properties of the forecasts obtained will be tested and the forecasting performance of alternative empirical models can then be analysed and compared. In addition, the resulting forecast errors, defined as actual minus forecast, can be decomposed into avoidable (systematic) and unavoidable (stochastic) components.

§7.3 Prediction and Evaluation of Estimated BVAR Forecasting Models for the European Economies

Once a BVAR or BVAR-EC forecasting model has been properly specified and estimated on some given set of data, the genuine 'out-of-sample' forecasts about future observations can be made by extrapolating the models beyond the period over which they were estimated. As well as being useful for policy formulations, ex ante forecasts also have the potential for evaluating model reliability. A forecast which is found to be way off target when actual data are available provides evidence which may lead to the revision of the model that provided the forecast (see Pindyck & Rubinfeld, *op. cit.*, pp.180-4).

7.3.1 Construction of Short-Term Successive Forecasts

In this work, the computation of successive 1-quarter ex-post forecasts from 1991Q1 onwards was done recursively using KF techniques to update the estimation with the passage of time first. Sequences of 1-quarter ex ante forecasts for the period one or two years ahead of the date of publication were then calculated based on the estimated BVAR and BVAR-EC models in both single- and multi-country context. The ex post and ex ante forecasted series of the variables being modelled with either a general (G) prior (narrow solid line) or a symmetric (S) prior (long dashed line) and the actual series (broad solid line) over the period 1991Q1-1998Q4 for the four principal European economies are appended and shown in Figs. 7.1 through 7.56.[2]

It would seem from these graphics that, for many variables, the fitted BVAR and BVAR-EC forecasting models have succeeded in capturing the overall trends of the actual series considered. What the forecasts fail to take into account is the 'lag-effects' tracking these variables at turning points of the business cycles.

[2]Again, only those figures for the UK are given here. The time plots of forecasts versus actuals for the other countries can be found in appendix F.

Some observed practical difficulties involved in the forecasting process need to be accentuated.

As can be seen from Figs. 7.1-7.4 for the logarithm of UK real GDP, while moving in the same directions as the original series, the forecasted series (made either in a single- or a multi-country context) cannot fully reflect the onset, depth and duration of the latest economic trough, namely 1991Q1-1992Q4. Similar problems have also arisen for Germany's real GNP over 1991Q1-1992Q4, and for the level of output in France and Italy over 1992Q2-1994Q2, as the period unfolded. Although this in itself may not be surprising, what is unusual is a significant consistent, or systematic, underprediction of UK output growth in cross-country forecasts over a two-year period from 1993Q1 to 1994Q4, and a sustained run of overpredictions in single-country BVAR-EC forecasts over the period from 1994Q4 onwards. The most likely explanation for the former is the lack of data on German real GNP, while the latter largely stems from the inclusion of the EC term much beyond the end of its estimation period.

Poor forecasts may happen when there are strong cyclical movements in the dependent variable that are inherently difficult to anticipate. In this case, even if a model has a good fit with statistically significant parameters, it may still forecast very weakly. A deterioration in forecasts may also occur when shifts in underlying behavioural relationships have not yet been fully incorporated into the model representation. If this poor performance appears to be due to an extraordinary event (such as dramatic oil-price developments or government policy changes) that are not accounted for by the model, another round of forecast has to take place before the model is proved unreliable. By contrast, good forecasts may happen when there is very little fluctuation in the dependent variable. In this case, even if a model has a relatively low R^2 and some insignificant regression coefficients, it may still predict quite easily. The forecasts for UK's M0 and CPI are notable examples.

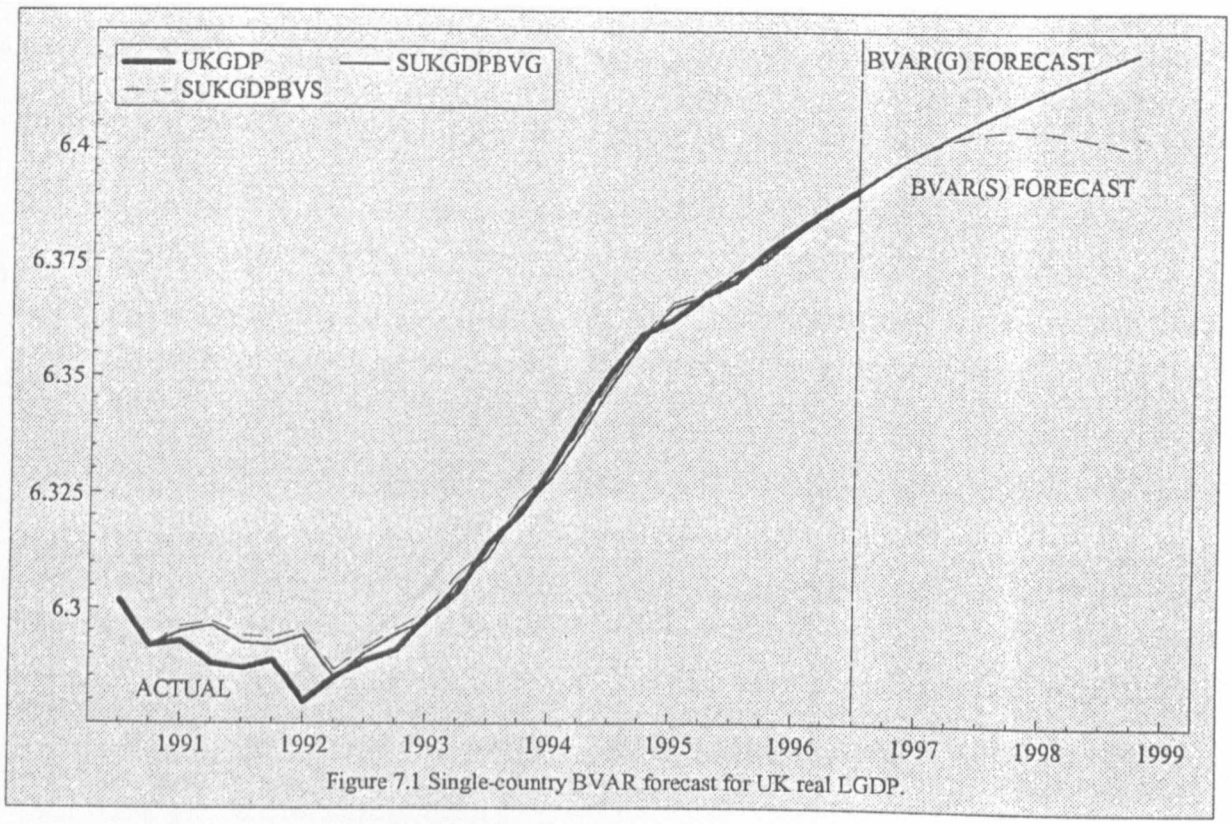

Figure 7.1 Single-country BVAR forecast for UK real LGDP.

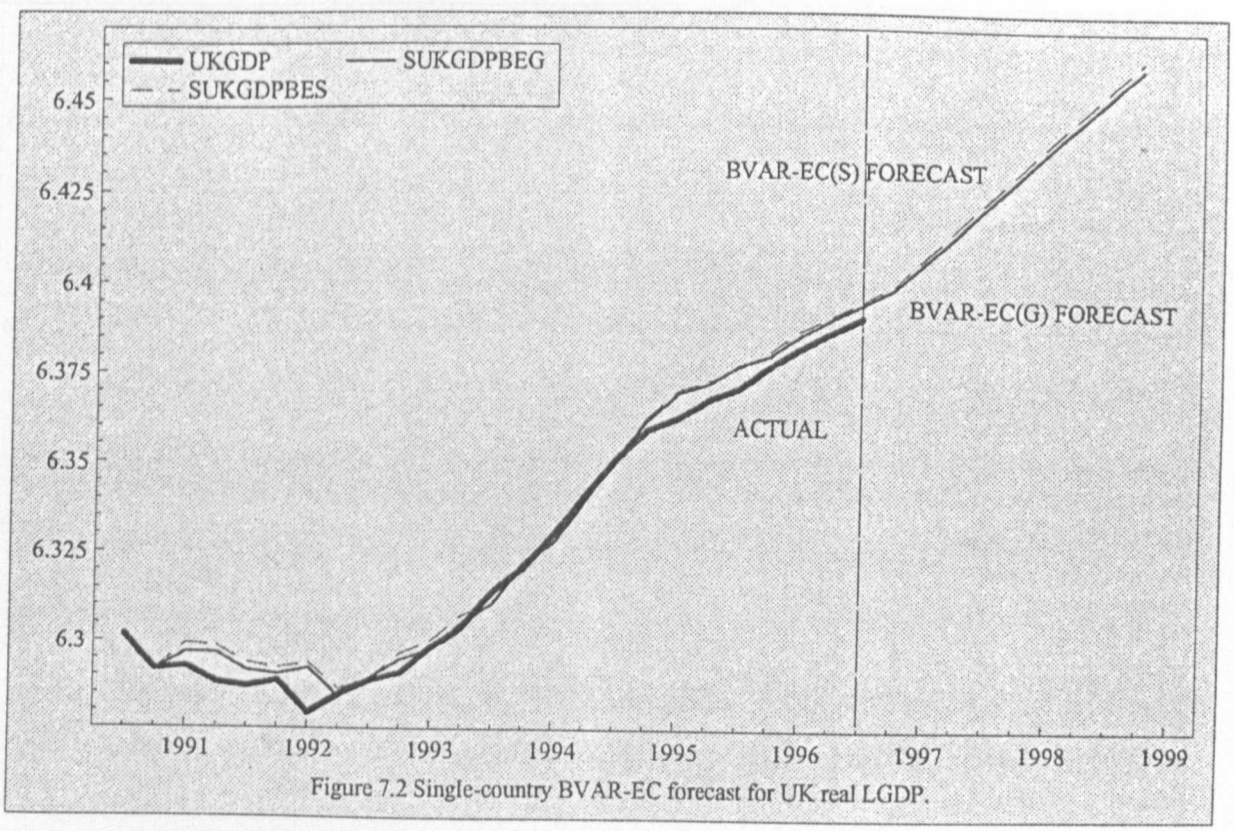

Figure 7.2 Single-country BVAR-EC forecast for UK real LGDP.

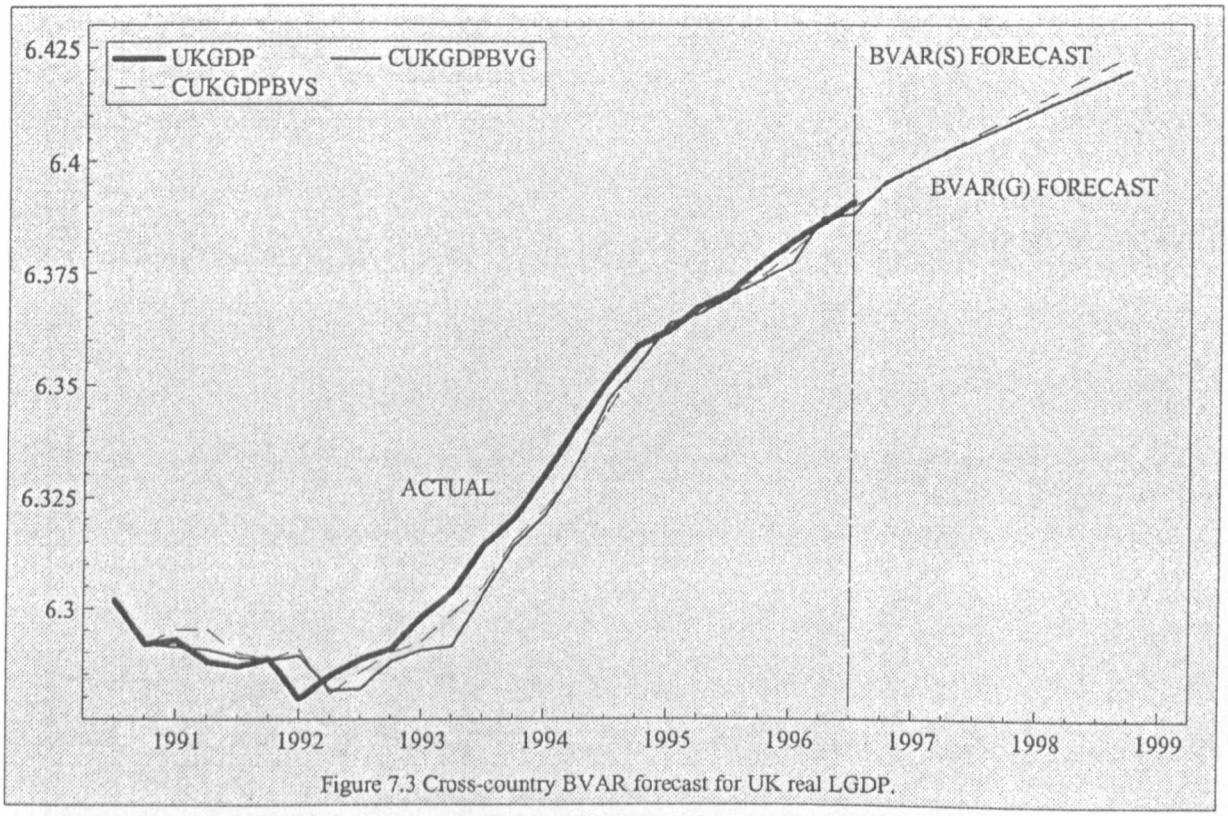

Figure 7.3 Cross-country BVAR forecast for UK real LGDP.

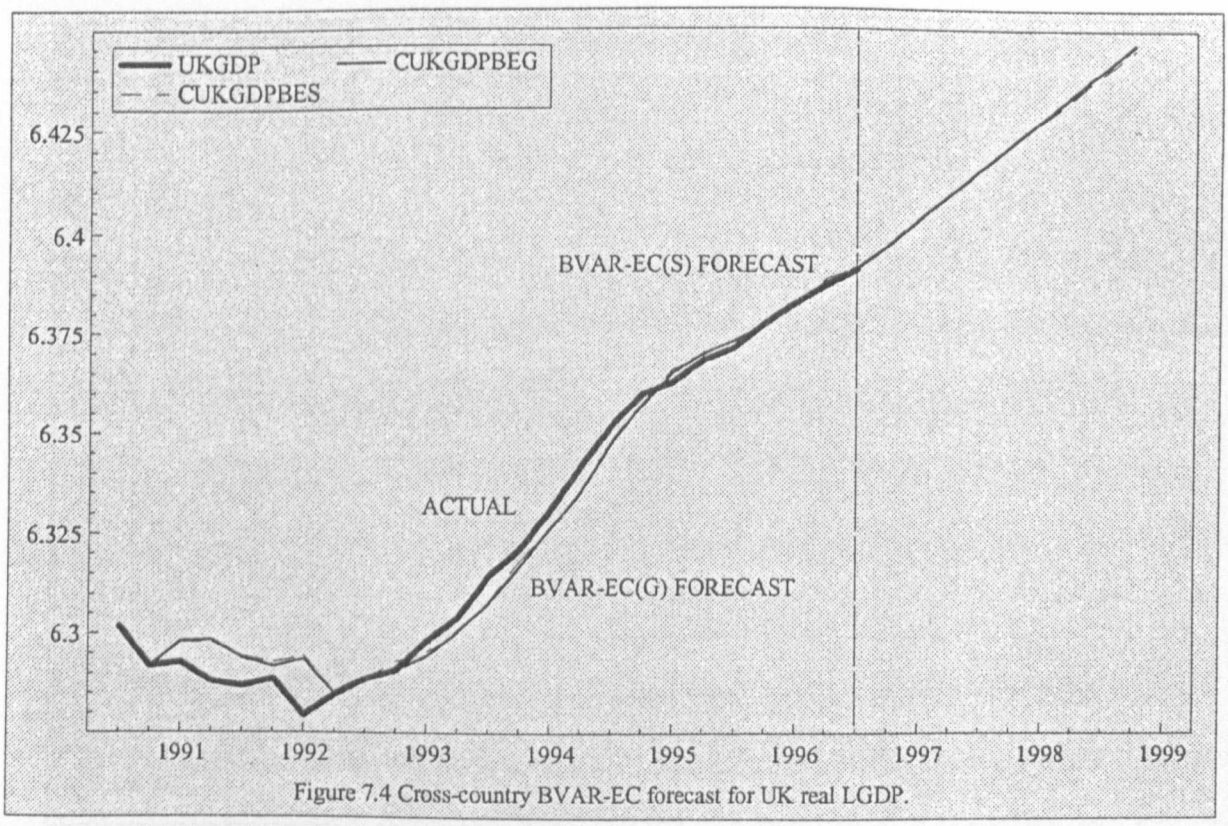

Figure 7.4 Cross-country BVAR-EC forecast for UK real LGDP.

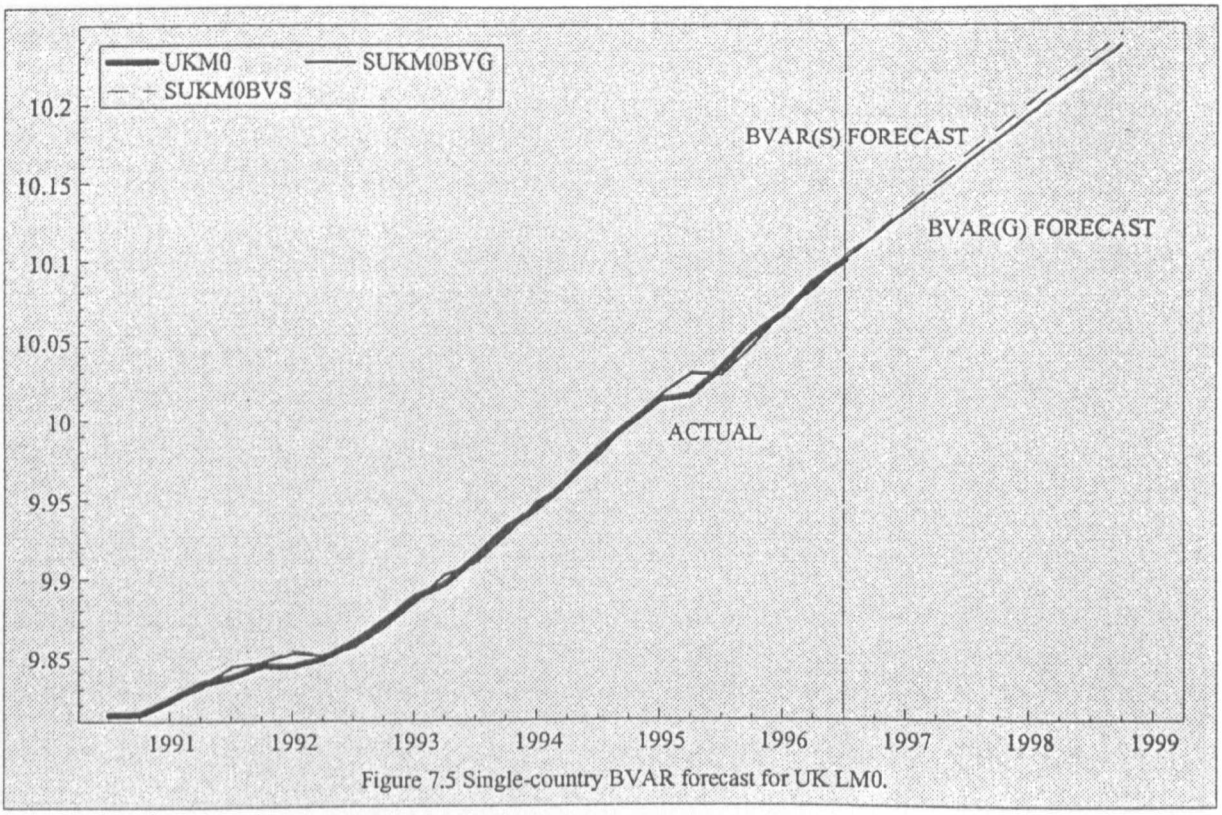

Figure 7.5 Single-country BVAR forecast for UK LM0.

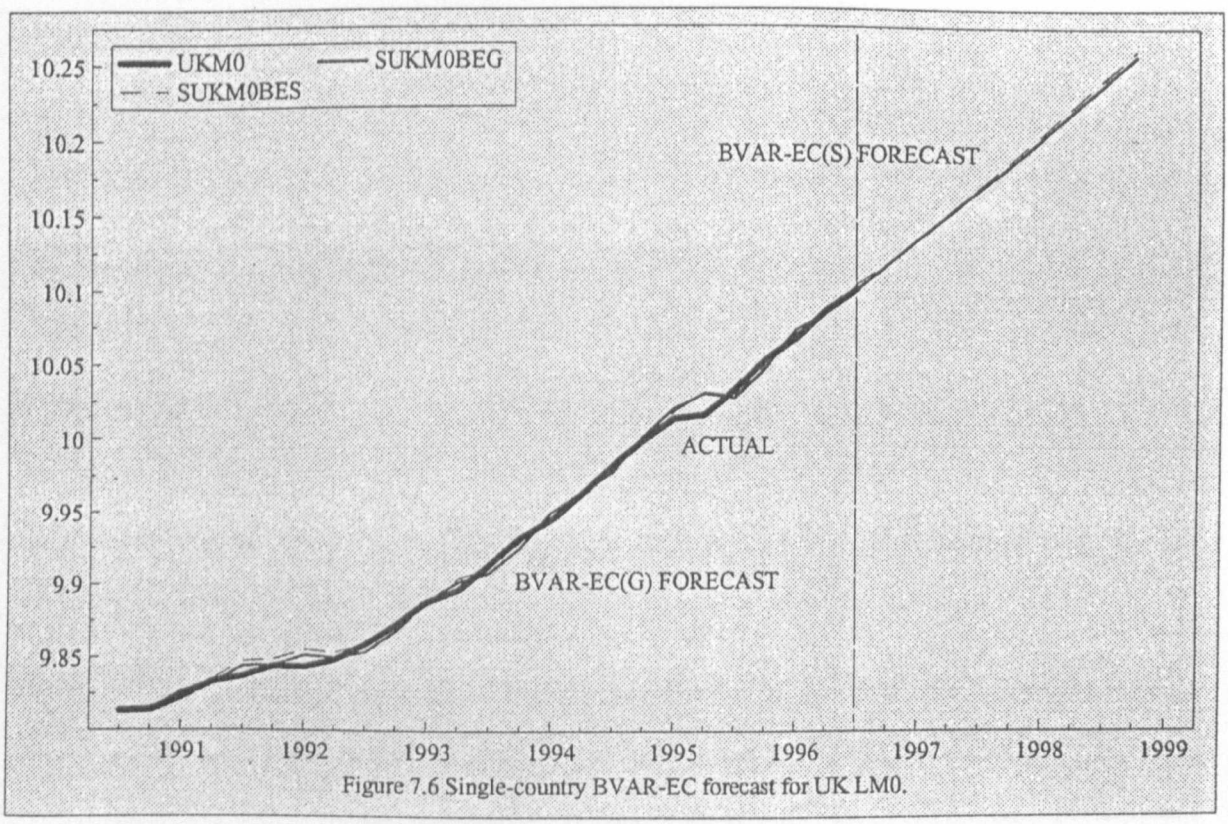

Figure 7.6 Single-country BVAR-EC forecast for UK LM0.

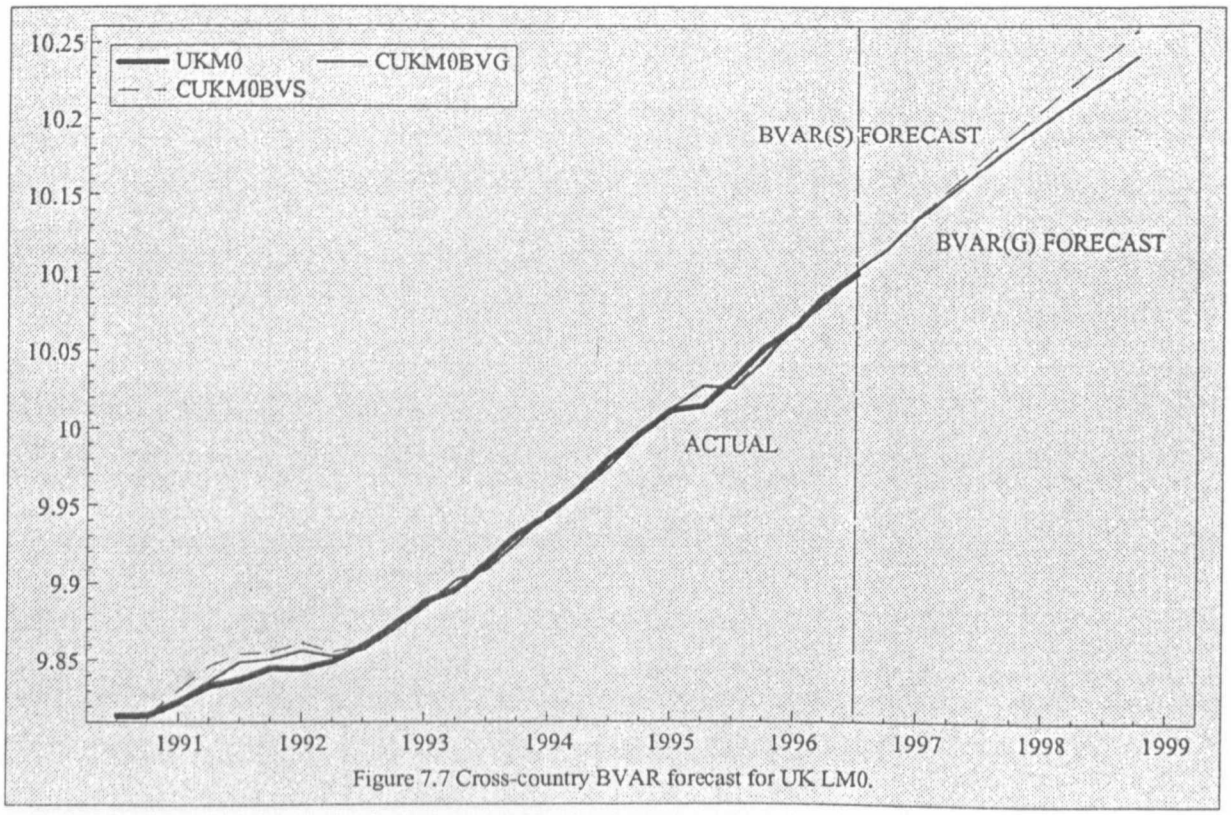

Figure 7.7 Cross-country BVAR forecast for UK LM0.

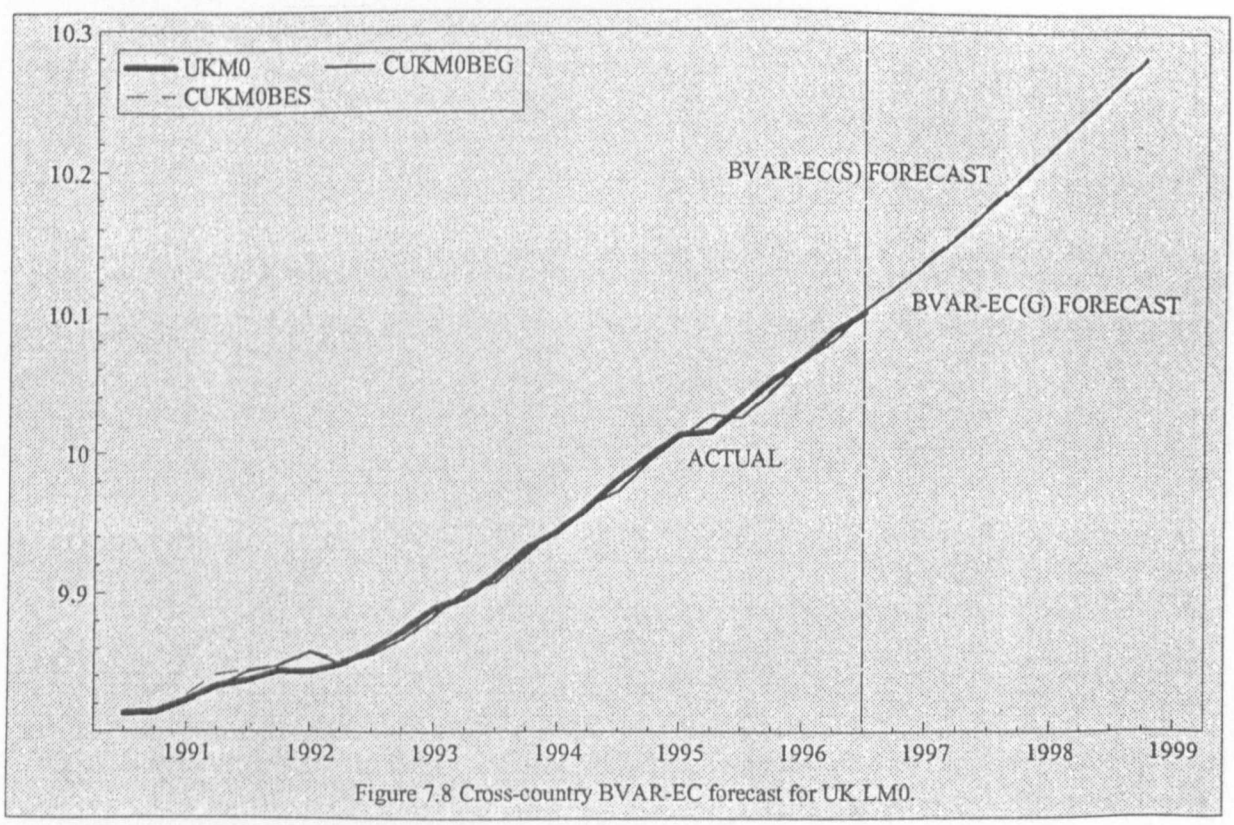

Figure 7.8 Cross-country BVAR-EC forecast for UK LM0.

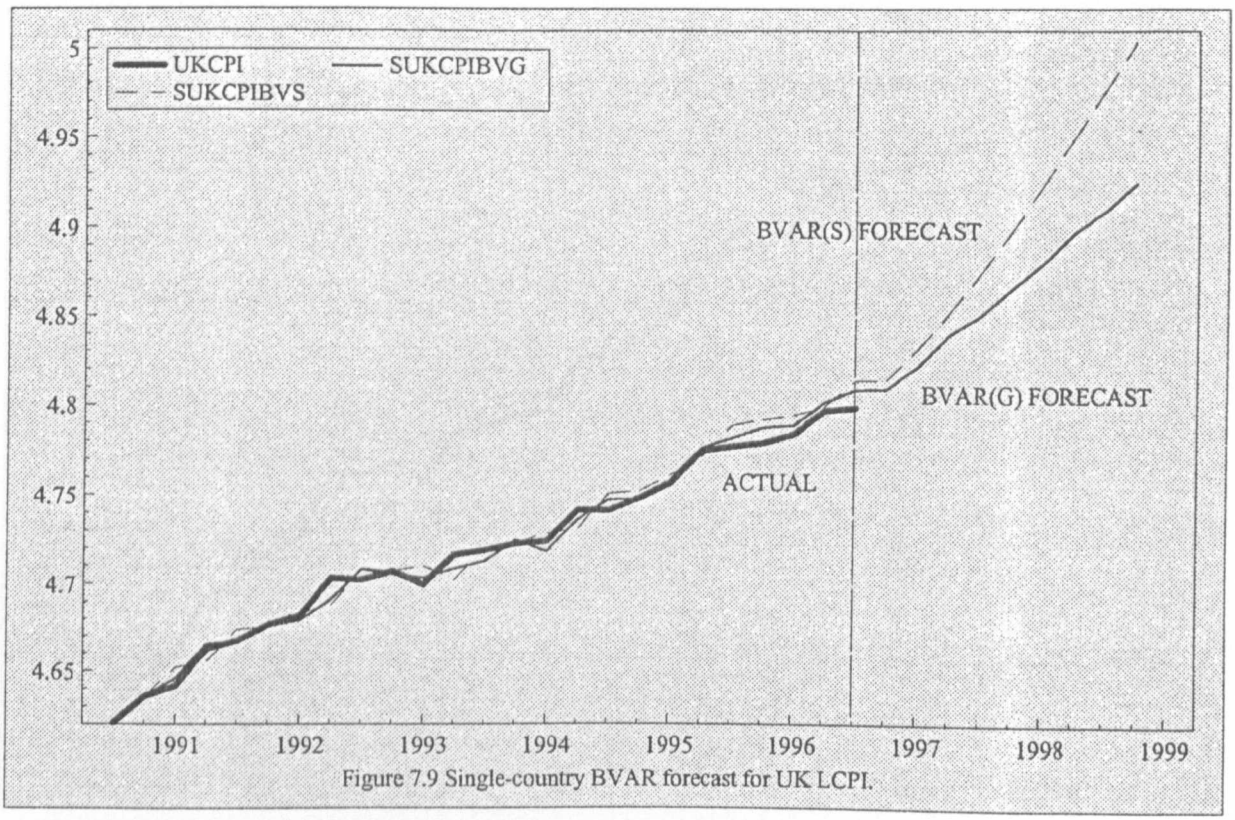

Figure 7.9 Single-country BVAR forecast for UK LCPI.

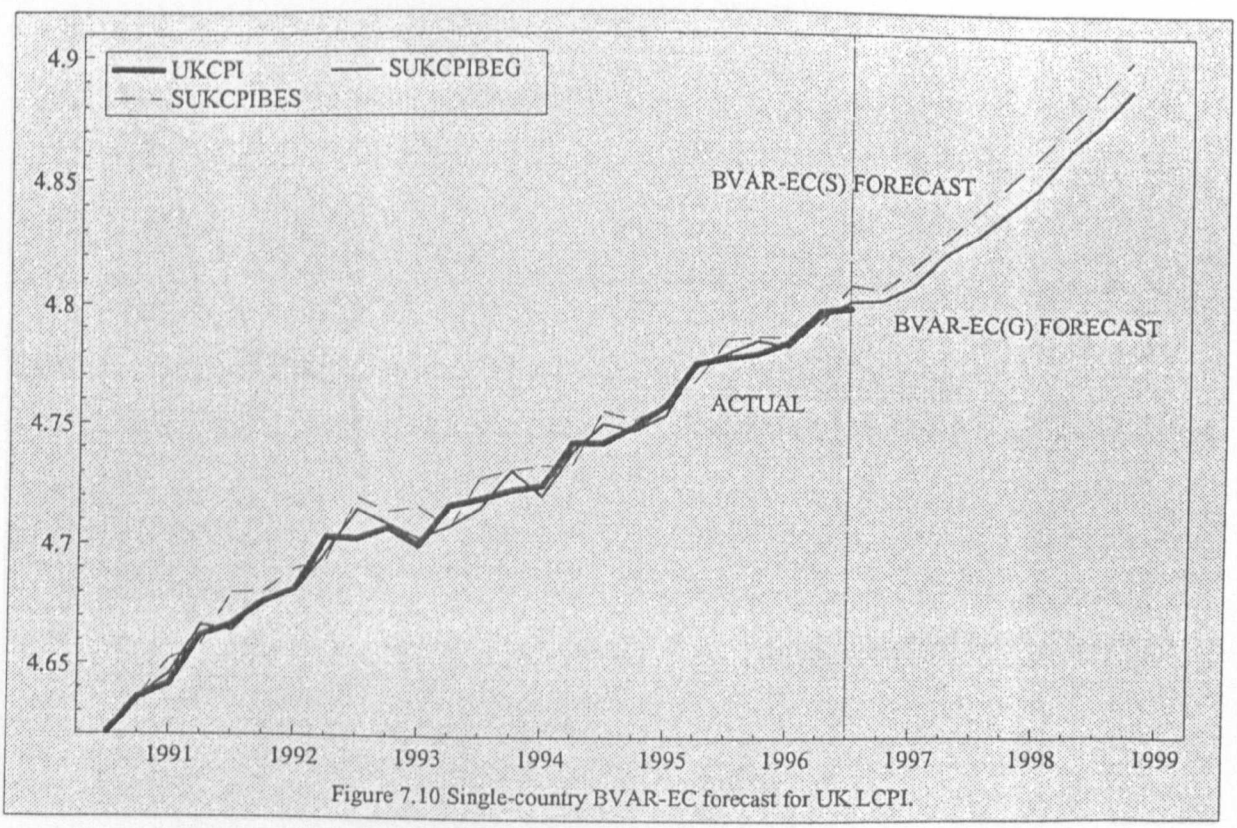

Figure 7.10 Single-country BVAR-EC forecast for UK LCPI.

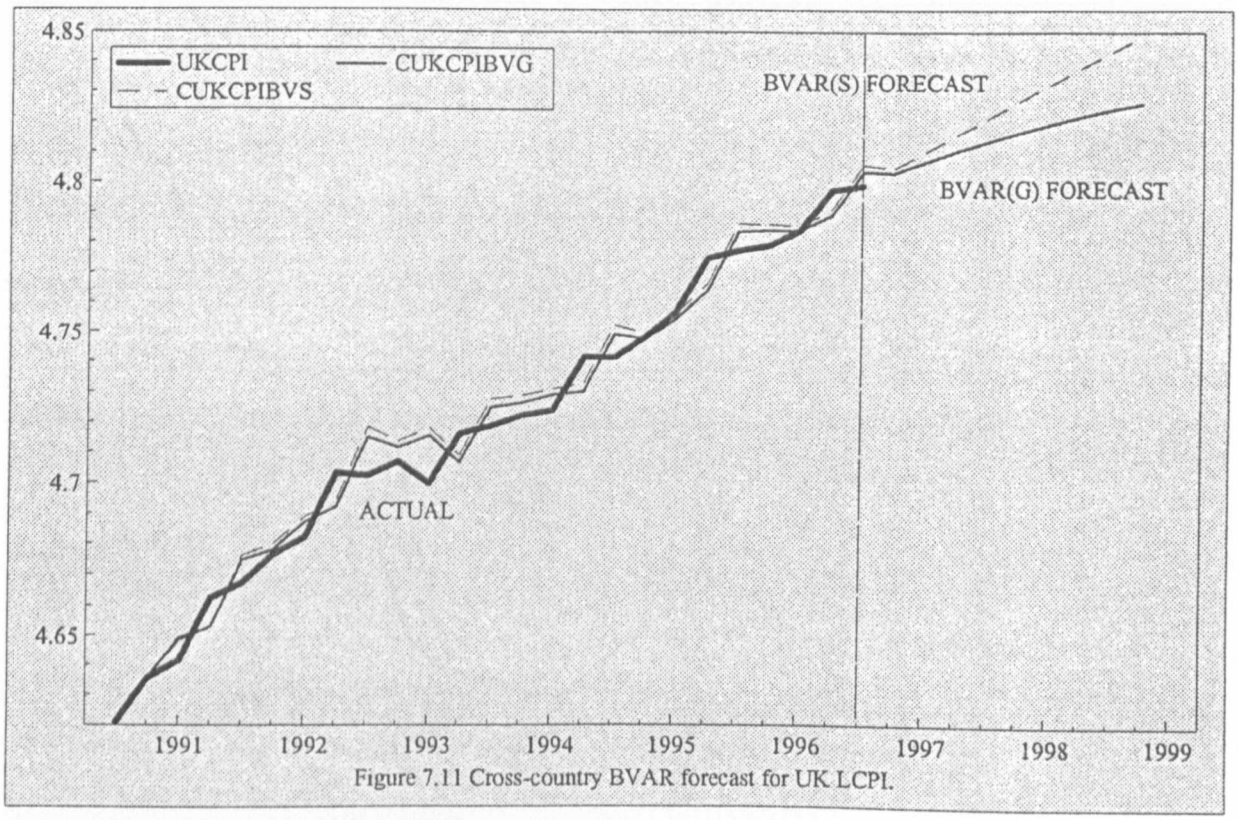

Figure 7.11 Cross-country BVAR forecast for UK LCPI.

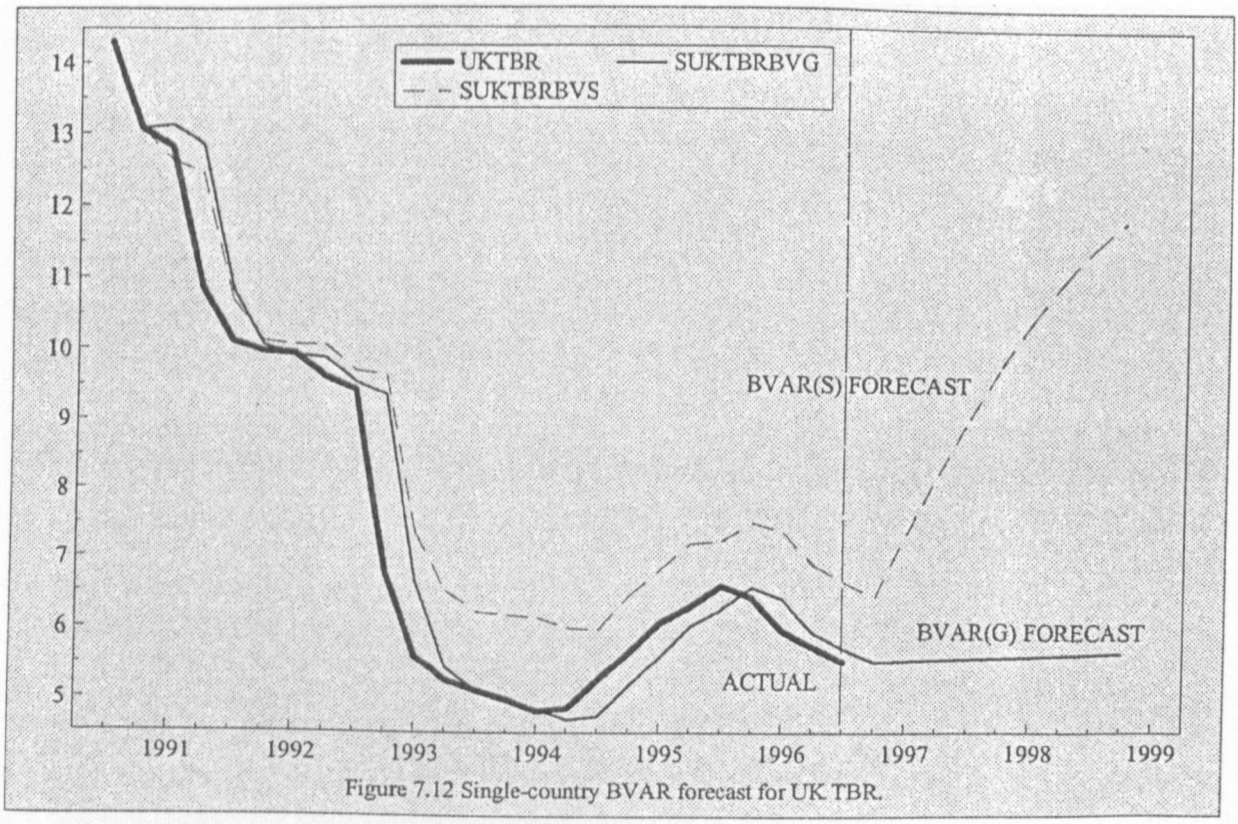

Figure 7.12 Single-country BVAR forecast for UK TBR.

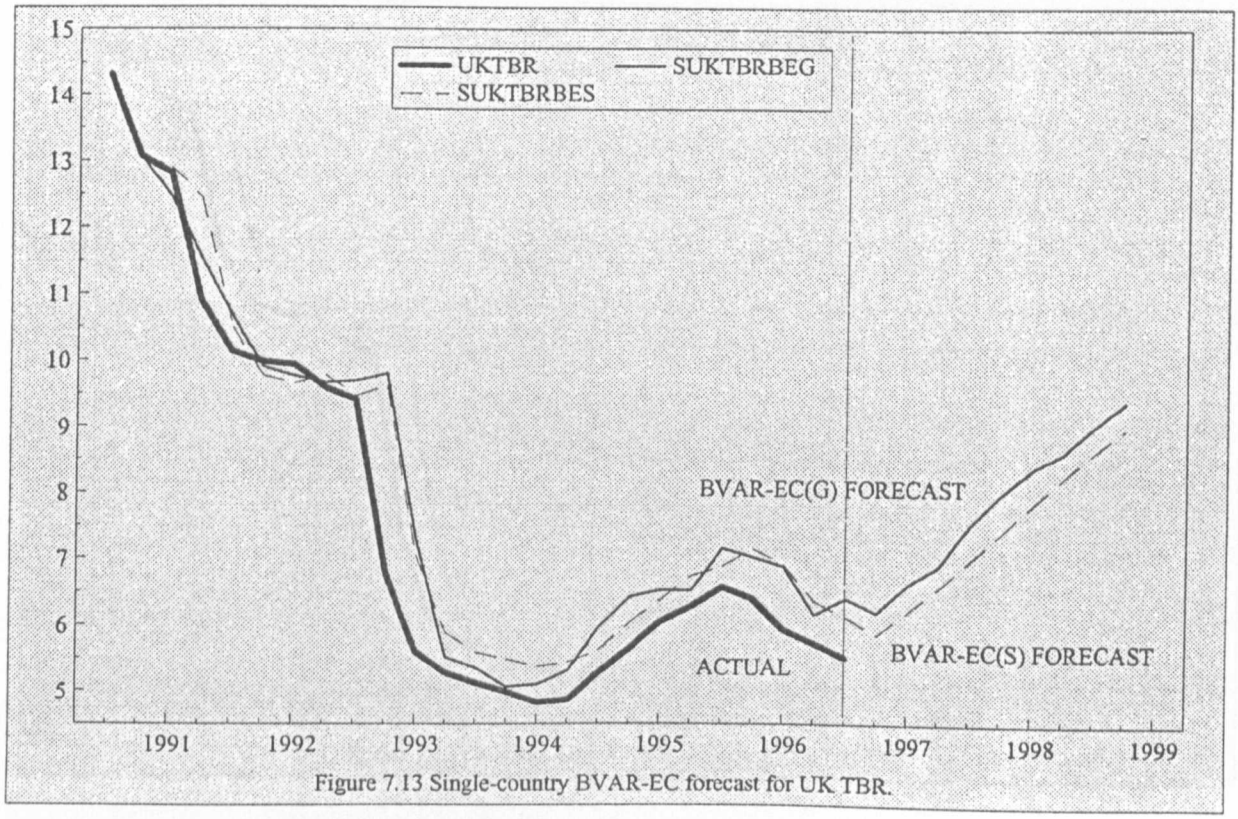

Figure 7.13 Single-country BVAR-EC forecast for UK TBR.

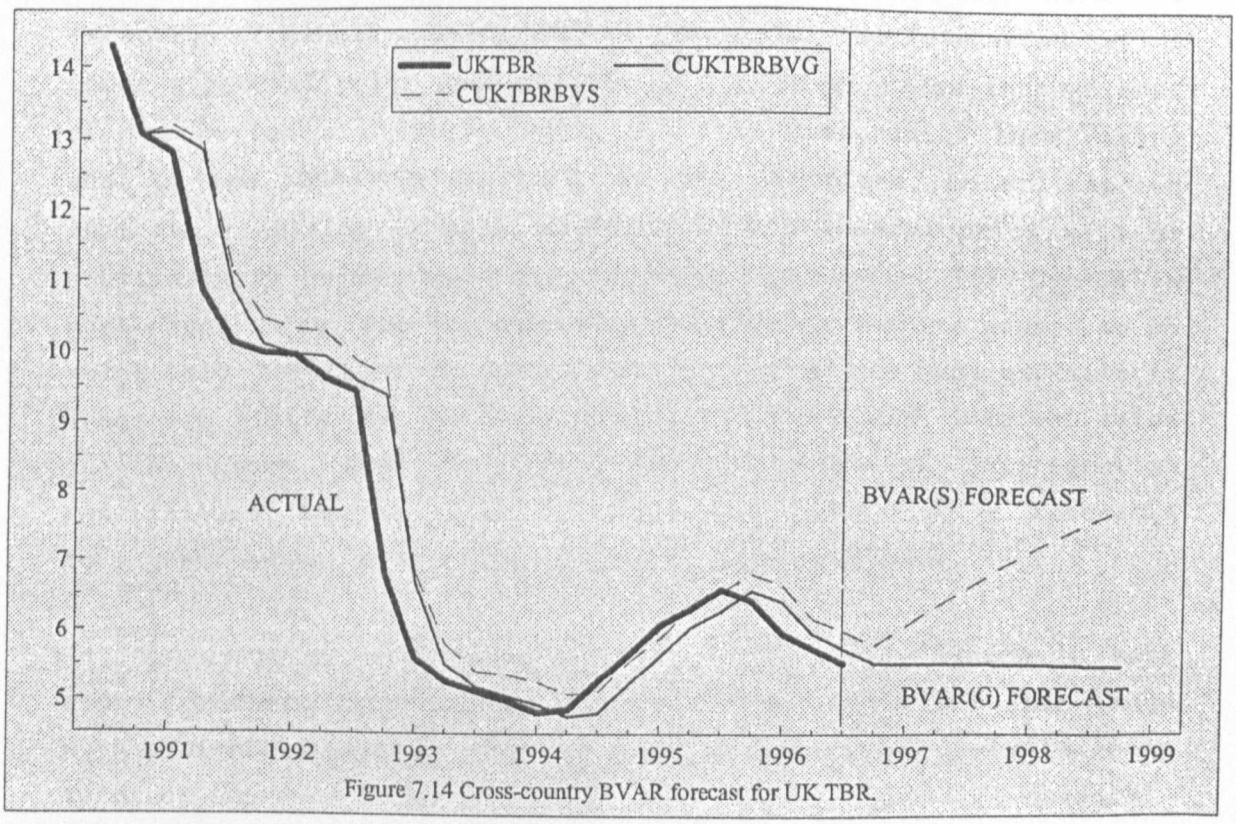

Figure 7.14 Cross-country BVAR forecast for UK TBR.

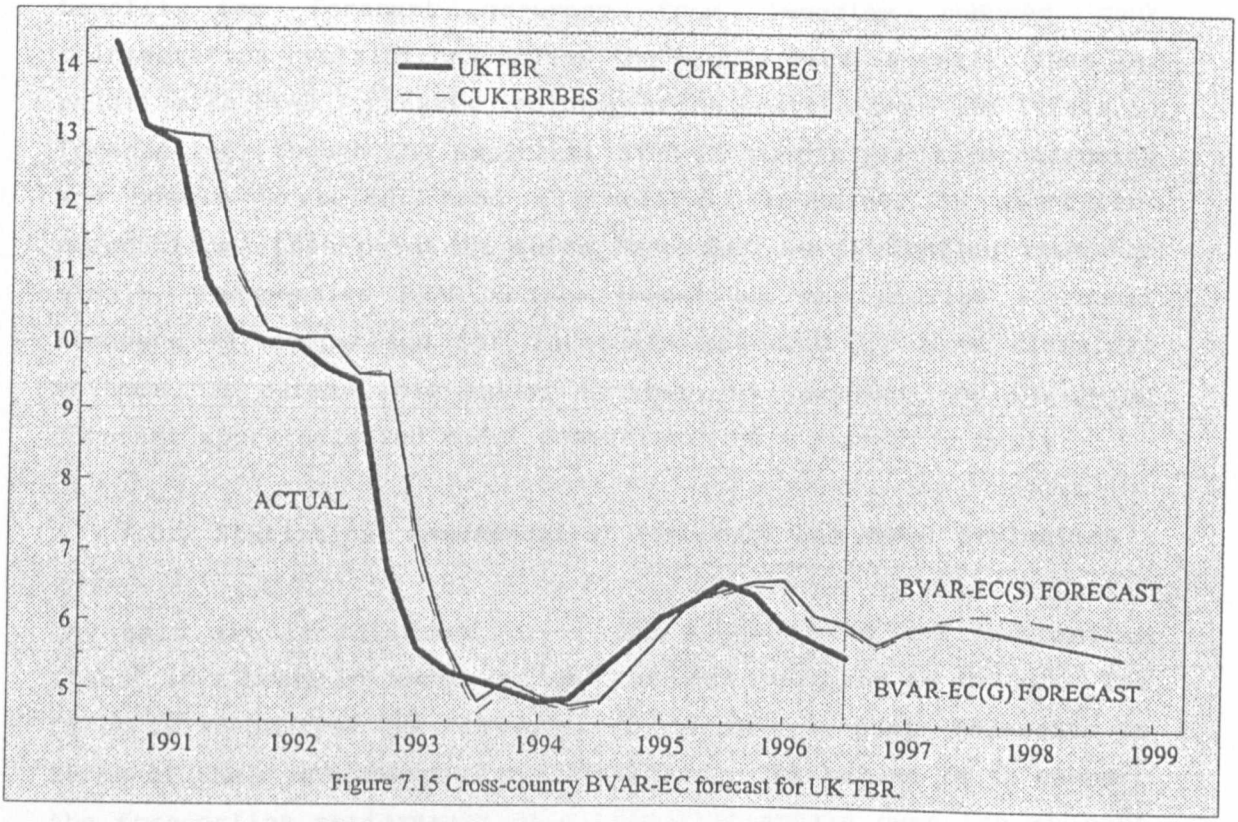

Figure 7.15 Cross-country BVAR-EC forecast for UK TBR.

Additionally, as shown in a group of Figs. 7.12-7.15, with the exception of single-country BVAR-EC forecasts, the forecasted series under a general prior underpredict UK's interest rates through most of the period of 1991Q1-1994Q1, but then overpredict them during most of the 1994-1995 quarters. It can, therefore, be anticipated that the resulting ex ante forecasts with a general prior will be substantially below the actual interest rate series for the UK in most cases. Also, the corresponding forecasting results appear to be relatively sensitive to variations in the prior hyperparameters. These are similar to the forecasts of the short-term interest rates in the other three countries. The inclusion of cointegration restrictions, however, seem to reduce the sensitivity of forecasts to a great extent across the values of hyperparameters.

As regards prior restrictions, it is found that although in most cases forecasts with a proper general prior are better than those with a simple symmetric one, the gains for models of modest size in terms of improved forecast accuracy are in many cases only marginal.

On the other hand, with respect to cointegration restrictions, it is found that as the forecast horizon lengthens, there is little benefit for forecast accuracy from imposing reduced rank cointegration restrictions unless the sample size is small. That is, relative to BVAR forecasts, cointegration restrictions add little to forecasting, though the empirical BVAR-EC models are time-consuming and costly to build. From a practical viewpoint, in short-term forecasting, this point is quite important as it implies that the simple, inexpensive BVAR models themselves may provide a robust standard of comparison for forecasts produced by more classical methods. To pursue this matter further, we consider the following forecast appraisals and model comparisons on a quarterly basis.

7.3.2 Statistical Assessment of Forecasts and Model Evaluation

The main aim in this section of the study is to use forecasting itself as a means of model evaluation. The intention is twofold. The first is to analyse the 'rationality' of past economic forecasts in terms of their unbiasedness and efficiency. The second is to compare the forecasting performance of alternative models over the ex post forecast horizons.

7.3.2a *Tests of the Efficiency and Unbiasedness Properties of the Forecasts*

An appropriate simple regression exercise allows us to examine whether the individual forecasts described above can be said to satisfy the minimum requirements expected of an efficient, or optimal, forecast.

(1) *A test of efficiency*

If Y_{jt} is taken to denote the actual outcome, and $_{t-i}\hat{Y}_{jt}$ the forecast for time t made at time t-i, one standard test of efficiency by Mincer and Zarnowitz (1969, p.9) involves a joint test of $\alpha_j=0$ and $\beta_j=1$ in the 'realisation-forecast' regression:

$$Y_{jt} = \alpha_j + \beta_j \, {}_{t-i}\hat{Y}_{jt} + \varepsilon_{jt}, \quad j = 1, \ldots, K. \tag{7.5}$$

Rejection of the null implies that the forecasts could be improved by knowledge of the α_j and β_j parameters and therefore provides evidence of inefficiency in the forecast. In particular, since

$$Y_{jt} \equiv {}_{t-i}\hat{Y}_{jt} + e_{jt}, \quad j = 1, \ldots, K,$$

the estimate of the slop coefficient (β_j) in the above regression only deviates from 1 if the forecast $({}_{t-i}\hat{Y}_{jt})$ and the forecast error (e_{jt}) are correlated. Such a correlation indicates that the forecast is not making efficient use of the given information and therefore it would be possible to improve the forecast by exploiting the correlation. For details, see Pain and Britton (1992, pp.81-93), and Wallis (1989, pp.44-6). A significant deviation of the estimates of α_j and β_j from 0 and 1, respectively, does not necessarily imply significant bias, for the efficiency hypothesis that $\alpha_j=0$ and $\beta_j=1$ is only a sufficient condition of unbiasedness, i.e.:

$$E(Y_{jt} - {}_{t-i}\hat{Y}_{jt}/I_{t-i-s}) = 0. \qquad (7.6)$$

Here, I_{t-i-s} is the information set available at the time of the forecast; s is the 'information lag' — the time between the latest available observed data and the time at which the forecast is prepared; and E denotes an expectations operator. A necessary and sufficient condition for the absence of bias is $\alpha_j=(1-\beta_j)E({}_{t-i}\hat{Y}_{jt})$, revealing that the forecast error has an expected value of zero.

Granger and Newbold (1986, p.284) raise a practical objection to the so-called 'efficiency' underlying the realisation-forecast regression. They argue that the hypothesis $\alpha_j=0$, $\beta_j=1$ constitutes only a necessary condition for forecast efficiency, since it neglects possible autocorrelation of the forecast error. Generally, the serial autocorrelation of errors indicates an inefficient forecast, as knowledge of past forecast errors for the variable in question can then improve current forecasts. Even in an optimal n-step-ahead forecast, such an autocorrelation may still appear but is of order n-1, not n; so it cannot be expected to improve the forecast [for further details about the test of efficiency, see Wallis (1989), pp.44-6].

An alternative approach proposed by Figlewski and Wachtel (1981) for efficiency is to test the joint hypothesis of $\alpha_j=\beta_j=0$ in a regression of the forecast error on the most recent error perceived at the time of the forecast:

$$Y_{jt} - {}_{t-i}\hat{Y}_{jt} = \alpha_j + \beta_j\left(Y_{j,t-i-1} - {}_{t-2i-1}\hat{Y}_{j,t-i-1}\right) + \varepsilon_{jt}, \quad (7.7)$$
$$j = 1, \ldots, K.$$

This means that efficient forecasts should make best use of the information contained in pass errors. In other words, forecasters always try to learn from their mistakes made in previous forecasts. See also Holden and Peel (1990), and Zellner (1986).

(2) *A test of unbiasedness*

A standard test for bias is thus to test the null that $\alpha_j=0$ in (7.8):

$$Y_{jt} - {}_{t-i}\hat{Y}_{jt} = \alpha_j + \varepsilon_{jt}, \quad j = 1, \ldots, K. \quad (7.8)$$

It is also useful to amend (7.7) and (7.8) so as to investigate whether the above forecasting results are sensitive to particular unanticipated exogenous shocks (such as an oil-price rise, financial shocks and major legislative changes). Such unforeseen events can generate large outliers in the observed forecast errors. The resulting equations augmented by dummy variables take the form:

$$Y_{jt} - {}_{t-i}\hat{Y}_{jt} = \alpha_j + \beta_j\left(Y_{j,t-i-1} - {}_{t-2i-1}\hat{Y}_{j,t-i-1}\right) + \delta_{j1}DUM(1) + \varepsilon_{jt}, \quad (7.9)$$

and

$$Y_{jt} - {}_{t-i}\hat{Y}_{jt} = \alpha_j + \rho_{j1}DUM(1) + \varepsilon_{jt}. \quad (7.10)$$

For simplicity, the empirical tests for the quarterly forecasts will be based on (7.7) and (7.8), and the corresponding testing results made since 1991Q1 are reported in tables 7.3 and 7.4.

Table 7.3: Test of efficiency hypothesis $\alpha_j=\beta_j=0$

$$(Y_{jt} - {}_{t-i}\hat{Y}_{jt} = \alpha_j + \beta_j (Y_{j,t-i-1} - {}_{t-2i-1}\hat{Y}_{j,t-i-1}) + \varepsilon_{jt})$$

Dep Var	SCM				MCM			
	BVAR		BVAR-EC		BVAR		BVAR-EC	
	α_j	β_j	α_j	β_j	α_j	β_j	α_j	β_j
UKLGDP	-.0009 (-.98)	.2801 (1.30)	-.0022 (-2.18)	.1881* (.86)	.0016 (1.42)	.5174* (2.71)	.0001 (.06)	.5956* (3.43)
UKLMO	-.0003 (-.26)	-.0421 (-.19)	-.0008 (-.63)	-.0395 (-.18)	-.0005 (-.38)	.1733 (.80)	.0009 (.67)	.0436 (.20)
UKLCPI	-.0004 (-.36)	.1820 (.78)	-.0005 (-.43)	-.2271 (-1.05)	-.0015 (-.96)	-.4186 (-2.08)	———	———
UKTBR	-.1554 (-.97)	.3883 (1.88)	-.5291 (-2.71)	.3157* (1.58)	-.1787 (-1.10)	.3851 (1.87)	-.2636 (-1.47)	.3654 (1.76)
GERLGNP	-.0090 (-2.43)	.1315* (.39)	-.0098 (-2.59)	.0545* (.16)	-.0029 (-.93)	.0294 (.08)	-.0123 (-2.72)	-.0208* (-.06)
GERLM1	-.0008 (-.20)	-.2638* (-2.55)	-.0082 (-2.72)	-.1125* (-1.42)	-.0002 (-.06)	-.1150 (-1.45)	-.0017 (-.57)	-.1174 (-1.45)
GERLCPI	.0016 (1.36)	.1110 (.50)	.0018 (1.48)	-.3943 (-1.92)	.0021 (2.10)	-.0995* (-.45)	———	———
GERTBR	-.1447 (-1.81)	.0052 (.02)	-.1516 (-1.89)	-.0125 (-.06)	.0151 (.20)	-.0827 (-.37)	-.3193 (-2.54)	.1046* (.47)
FRALGDP	-.0006 (-.46)	.1244 (.58)	-.0022 (-1.37)	.2036 (.93)	-.0005 (-.35)	.2017 (.92)	-.0026 (-1.59)	.0234 (.10)
FRALM1	-.0006 (-.17)	.1118 (.56)	.0009 (.22)	.1540 (.77)	-.0034 (-.98)	.1201 (.63)	-.0142 (-3.04)	.1156* (.60)
FRALCPI	.0000 (.05)	-.1955 (-.78)	.0002 (.22)	.0673 (.28)	-.0014 (-1.89)	.0483 (.20)	———	———
FRATBR	-.1106 (-.50)	-.0239 (-.11)	-.2048 (-.91)	-.0251 (-.11)	-.4421 (-1.97)	-.0691 (-.31)	-1.0535 (-4.02)	-.2486* (-1.15)
ITALGDP	-.0009 (-.74)	-.1310 (-.59)	———	———	-.0046 (-3.75)	-.5507* (-2.97)	-.0107 (-5.83)	-.5319* (-2.75)
ITALM1	-.0038 (-.96)	-.0953 (-.42)	———	———	-.0072 (-1.79)	-.0543 (-.25)	-.0186 (-3.49)	-.0881* (-.40)
ITALCPI	-.0001 (-.15)	-.0177 (-.07)	———	———	.0001 (.16)	.2873 (1.10)	———	———
ITATBR	-.0394 (-.17)	.3250 (1.54)	———	———	-.0969 (-.44)	.2261 (1.04)	-.1586 (-.65)	.3098 (1.47)

Note: (i) SCM stands for single-country model and MCM for multi-country model;

(ii) The values within (·) are t-ratios;

(iii) * indicates inefficient forecast at 5% significance level.

Table 7.4: Test of unbiasedness property $\alpha_j=0$ $\left(Y_{jt}-{}_{t-i}\hat{Y}_{jt}=\alpha_j+\varepsilon_{jt}\right)$

Dep Var	SCM		MCM	
	BVAR	BVAR-EC	BVAR	BVAR-EC
	α_j	α_j	α_j	α_j
UKLGDP	-.0013 (-1.57)	-.0028* (-3.60)	.0032* (3.10)	-.0004 (-.41)
UKLM0	-.0002 (-.24)	-.0009 (-.76)	-.0006 (-.54)	.0008 (.69)
UKLCPI	-.0007 (-.57)	-.0005 (-.51)	-.0013 (-.79)	———
UKTBR	-.2587 (-1.69)	-.6990* (-4.80)	-.2922 (-1.93)	-.3945* (-2.45)
GERLGNP	-.0075* (-2.20)	-.0080* (-2.50)	-.0009 (-.27)	-.0102* (-3.32)
GERLM1	.0048 (.57)	-.0008 (-.10)	.0061 (.81)	.0046 (.63)
GERLCPI	.0019 (1.88)	.0014 (1.10)	.0019* (2.11)	———
GERTBR	-.1401 (-2.03)	-.1494* (-2.22)	.0171 (.25)	-.3419* (-3.56)
FRALGDP	-.0011 (-.87)	-.0031* (-2.29)	-.0010 (-.74)	-.0029* (-2.12)
FRALM1	-.0025 (-.69)	-.0010 (-.23)	-.0057 (-1.65)	-.0178* (-5.35)
FRALCPI	.0000 (.01)	.0002 (.23)	-.0016* (-2.48)	———
FRATBR	-.1066 (-.52)	-.1990 (-.97)	-.4092* (-2.14)	-.8293* (-4.49)
ITALGDP	-.0007 (-.64)	———	-.0028* (-2.22)	-.0067* (-4.99)
ITALM1	-.0036 (-1.00)	———	-.0070 (-2.01)	-.0172* (-4.71)
ITALCPI	.0000 (.03)	———	.0003 (.47)	———
ITATBR	-.0255 (-.11)	———	-.1403 (-.67)	-.1952 (-.83)

Note: (i) SCM stands for single-country model and MCM for multi-country model;
(ii) The values within (·) are t-ratios;
(iii) * indicates significant bias at 5% level.

Of the 56 cases considered in Table 7.3, the efficiency hypothesis is rejected at the 5% significance level in 16 cases. But 19 of the 56 cases are found to be biased according to Table 7.4, with many of the inefficient forecasts also being biased. Comparing the two tables, it can be seen that in most cases the inefficiencies in the forecasts associated with a significant intercept term (α_j) will simply confirm the biases, and that many such biases appear to be upward, implying a systematic overprediction over the forecast period as a whole.

The results are consistent with earlier studies in showing that the observed forecasting difficulties arise largely from GDP/GNP growth

and short-term interest rates across four selected variables, which account for nearly 70% and 80% of rejections in the respective hypotheses tests of the forecasts. Especially for forecasts of Germany's real GNP, only one of the four cases, namely the cross-country BVAR forecasts (one-step-ahead) over the period 1991Q1-1992Q4, fail to reject the null of efficiency and of unbiasedness. However, as with the previous studies, the evidence of efficiency appears greatest in the forecasts of single-country BVAR forecasting models over the period from 1991Q1 onwards, with the null only being rejected for Germany's GNP and M1 at conventional levels. These results can be further confirmed through the following forecast comparisons across alternative models.

7.3.2b *Comparison of Alternative Forecasting Models*

The forecasting literature has a long history of assessing the relative accuracies in predictions of alternative models using appropriate statistical tests, and of taking these assessments as an important input into the forecasting process, since without comparison we could never say which one is 'good' and which one is 'poor'. A useful, but controversial, principle is that if one group consistently outperforms another, then the poorer of the two cannot sensibly be used for policy purposes, which makes forecast comparisons important.

For simplicity, only the one-quarter-ahead forecasting performance of the BVAR versions is systematically compared with that of their BVAR-EC counterparts for Europe's major four economies, based upon the Theil U statistics reported in Tables 7.5 through 7.8. Since the two groups of forecasters 'breath the same air' with exactly the same vector contents and forecast horizons, the tedious problem of the 'alignment' of base-line information assumptions embodied in different systems can be avoided.

The comparison of Theil U statistics in Table 7.5 demonstrates a relative superiority of the single-country BVAR models over the other specifications for the UK economy. For the one-quarter forecasts of the four macroeconomic aggregates of interest, the

single-country BVAR forecasts have the smallest Theil U values in 3 out of 4 cases and all of these values are less than unity. Additionally, the root mean squared errors (RMSEs) are well above the mean absolute errors (MAEs) in all cases.

Table 7.5: One-step-ahead forecast comparison for UK economy: BVAR and BVAR-EC

Dep Var	Model		Criteria			
			ME	MAE	RMSE	Theil U
UKLGDP	SCM-UK	BVAR	-0.0013	0.0028	0.0042	0.6573*
		BVAR-EC	-0.0052	0.0055	0.0069	0.7528
	MCM-LGDP	BVAR	0.0000	0.0036	0.0046	1.0429
		BVAR-EC	-0.0078	0.0078	0.0080	0.8897
UKLMO	SCM-UK	BVAR	-0.0002	0.0040	0.0049	0.3582*
		BVAR-EC	-0.0031	0.0036	0.0047	0.7873
	MCM-LMON	BVAR	-0.0006	0.0046	0.0057	0.4082
		BVAR-EC	-0.0022	0.0030	0.0037	0.6866
UKLCPI	SCM-UK	BVAR	-0.0007	0.0046	0.0054	0.5339
		BVAR-EC	0.0001	0.0031	0.0036	0.3065*
	MCM-LCPI	BVAR	-0.0013	0.0068	0.0079	0.7723
UKTBR	SCM-UK	BVAR	-0.2587	0.4575	0.7645	0.9792*
		BVAR-EC	-0.3083	0.5101	0.7825	0.9937
	MCM-TBR	BVAR	-0.2922	0.4659	0.7677	0.9833
		BVAR-EC	-0.4865	0.4865	0.8344	1.4800

Note: (i) SCM stands for single-country model; MCM for multi-country model; ME for mean error; MAE for mean absolute error; and RMSE for root mean squared error;
(ii) * indicates the ex-post forecast with the lowest Theil U statistic.

The evidence that the BVAR model is substantially more accurate than its BVAR-EC counterpart may also be shown from Table 7.6 in the forecasts for the German economy. Although the forecast results in this case cannot be matched with those in the previous one, the relative pattern of forecast performance is not reversed. The resulting Theil U statistics indicate the comparative advantage of the single-country, and of the cross-country, BVAR models over the

other versions in a quarter, and a half, respectively, of the number of cases examined.

Table 7.6: One-step-ahead forecast comparison for GERMAN economy: BVAR and BVAR-EC

Dep Var	Model		Criteria			
			ME	MAE	RMSE	Theil U
GELGNP	SCM-GERMANY	BVAR	-0.0075	0.0104	0.0117	1.3272
		BVAR-EC	-0.0104	0.0120	0.0134	1.0600
	MCM-LGDP	BVAR	-0.0009	0.0071	0.0085	0.9638*
		BVAR-EC	-0.0086	0.0128	0.0145	1.2226
GELM1	SCM-GERMANY	BVAR	0.0147	0.0408	0.0645	0.9518
		BVAR-EC	-0.0024	0.0149	0.0155	1.3343
	MCM-LMON	BVAR	0.0061	0.0172	0.0363	0.8421
		BVAR-EC	0.0310	0.0493	0.0811	0.6934*
GELCPI	SCM-GERMANY	BVAR	0.0041	0.0044	0.0051	0.5183*
		BVAR-EC	0.0031	0.0045	0.0059	1.2878
	MCM-LCPI	BVAR	0.0019	0.0035	0.0045	0.5235
GETBR	SCM-GERMANY	BVAR	-0.0605	0.2389	0.3758	0.9081
		BVAR-EC	-0.3989	0.3989	0.6786	0.9831
	MCM-TBR	BVAR	0.0171	0.2591	0.3284	0.8339*
		BVAR-EC	-0.1366	0.1366	0.1874	0.8509

Note: (i) SCM stands for single-country model; MCM for multi-country model; ME for mean error; MAE for mean absolute error; and RMSE for Root mean squared error;
(ii) * indicates the ex-post forecast with the lowest Theil U statistic.

A perusal of the Theil U statistics summarized in Table 7.7 finds that the BVAR models outperform their BVAR-EC counterparts in most cases for the French economy. The single-country BVAR models have a superior forecasting performance to the others in 3 out of 4 instances, although there is one case in which the Theil U statistic exceeds unity.

Table 7.7: One-step-ahead forecast comparison for FRENCH economy: BVAR and BVAR-EC

Dep Var	Model		Criteria			
			ME	MAE	RMSE	Theil U
FRLGDP	SCM-FRANCE	BVAR	-0.0011	0.0047	0.0062	0.8949*
		BVAR-EC	-0.0041	0.0041	0.0052	1.3308
	MCM-LGDP	BVAR	-0.0027	0.0038	0.0045	1.0719
		BVAR-EC	-0.0068	0.0068	0.0074	1.5164
FRLM1	SCM-FRANCE	BVAR	-0.0025	0.0123	0.0168	1.0725*
		BVAR-EC	-0.0199	0.0199	0.0231	1.0831
	MCM-LMON	BVAR	-0.0057	0.0133	0.0176	1.1409
		BVAR-EC	-0.0349	0.0349	0.0371	1.2890
FRLCPI	SCM-FRANCE	BVAR	0.0000	0.0019	0.0026	0.4587*
		BVAR-EC	-0.0012	0.0023	0.0030	0.5223
	MCM-LCPI	BVAR	-0.0016	0.0025	0.0033	0.5850
FRTBR	SCM-FRANCE	BVAR	-0.1066	0.6083	0.9742	0.9810
		BVAR-EC	-0.0919	0.5401	0.8626	0.8987*
	MCM-TBR	BVAR	-0.4092	0.6703	0.9861	0.9930
		BVAR-EC	-0.3736	0.7512	1.0681	1.0934

Note: (i) SCM stands for single-country model; MCM for multi-country model; ME for mean error; MAE for mean absolute error; and RMSE for root mean squared error;
(ii) * indicates the ex-post forecast with the lowest Theil U statistic.

Finally, as shown in Table 7.8, the BVAR models again prove to be capable of producing forecasts comparable in quality for the Italian economy. According to the table, the results of single-country BVAR forecasts would still appear to be fairly encouraging and the corresponding U values reported here are all satisfactorily below unity.

Table 7.8: One-step-ahead forecast comparison for ITALIAN economy: BVAR and BVAR-EC

Dep Var	Model		Criteria			
			ME	MAE	RMSE	Theil U
ITLGDP	SCM-ITALY	BVAR	-0.0007	0.0040	0.0055	0.8206*
	MCM-LGDP	BVAR	-0.0025	0.0031	0.0050	1.0168
		BVAR-EC	-0.0051	0.0051	0.0064	1.3402
ITLM1	SCM-ITALY	BVAR	-0.0036	0.0139	0.0172	0.8417*
	MCM-LMON	BVAR	-0.0070	0.0150	0.0181	0.8980
		BVAR-EC	-0.0005	0.0131	0.0159	0.8676
ITLCPI	SCM-ITALY	BVAR	0.0000	0.0025	0.0033	0.2744
	MCM-LCPI	BVAR	0.0003	0.0021	0.0031	0.2580*
ITTBR	SCM-ITALY	BVAR	-0.0255	0.7981	1.0886	0.9999
	MCM-TBR	BVAR	-0.1403	0.7207	0.9884	0.9078
		BVAR-EC	0.4064	0.6367	0.7553	0.8000*

Note: (i) SCM stands for single-country model; MCM for multi-country model; ME for mean error; MAE for mean absolute error; and RMSE for root mean squared error;
(ii) * indicates the ex-post forecast with the lowest Theil U statistic.

Overall the results are mixed. It may be concluded from this particular application that the single-country BVAR forecasts one-quarter-ahead are, on average, marginally superior for the four leading European countries across the alternative forecasting models under study. Of the 16 comparisons, the single-country BVAR forecasts are more accurate in 9 cases. Perhaps, the least accurate forecasts generated by both BVAR and BVAR-EC models are those associated with the interest rate on Treasury bills, where the reported Theil U values generally appear to be the largest throughout the experiment.

§7.4 Conclusion

In this chapter, the empirical BVAR forecasting models are estimated and employed to generate one-quarter-ahead ex post and ex ante forecasts of some key macroeconomic variables for the four major European economies with the aid of KF estimation technique. The full results obtained from the mixed-estimated BVAR and BVAR-EC models are illustrated with charts over the forecast horizon, 1991Q1-1998Q4, in both specific- and cross-country context and as reported in Appendix F. The assessments of the efficiency and bias of the forecasts as well as the comparisons of the Theil U statistics indicate that in overall terms the single-country BVAR models emerge as providing a high standard of forecasting, especially for small sample sizes and long forecast horizons.

However, this conclusion has to be treated as tentative. An improvement for forecast accuracy could be gained occasionally from imposing cointegrated combinations of variables within or across the countries. In fact, such cross-model comparisons cannot be used to rank models because, as we observed, under an appropriate forecasting criterion (e.g. Theil's U statistic), no one forecast set 'dominates' another for all variables over all time horizons. It would, therefore, be more fruitful to regard several different approaches less as rivals than as complementary tools that can shed different kinds of light from different viewpoints on our projections of what the potential rule for each of the series would be like in the future. See Granger (ed) (1991), pp.1-23.

CHAPTER 8

SUMMARY, CONCLUSIONS AND POTENTIAL DEVELOPMENTS

§8.1 Summary

Recent forecast failures in the world economy clearly highlight the importance of exploiting the sources of forecast accuracy, both for their substantive implications and for their implications as to how we should improve our methodologies in order to provide the best strategy for macroeconometric modelling. It is now generally recognized that macroeconometric models are important tools in the hands of analysts. They are used, *inter alia*, for explaining economic phenomena, making forecasts and assessing policy changes, and their usefulness is unlikely to be superceded by other methods in the foreseeable future.

However, much work still needs to be done on alternative research strategies. The economic theory-oriented approach related to the Lucas critique and the rational expectations hypothesis (see, for example, Lucas and Sargent, 1981) has criticized the modelling of expectations and casts considerable doubt upon the invariance of so-called structural parameters. The time series-oriented criticism has questioned the existence of legitimate prior theoretical knowledge and emphasized the significance of the joint temporal structure of the observed data on relevant variables (Sims, 1980a, 1982). The Bayesian-oriented school of thought, for its part, has criticized the *ad hoc* nature of traditional econometrics and prescribed a more formal treatment using appropriate techniques (Litterman, 1980, 1986a).

The explicitly Bayesian forecasting procedure with stochastic priors provides key insights into the process of macroeconometric modelling. It allows the problem of overfitting to be ameliorated,

and creates useful and natural vehicles for working with the available data (Shoesmith, 1990, 1992). In this study, a systematic, practical six-step procedure has been developed for the formulation, estimation and selection of appropriate BVAR and BVAR-EC forecasting models. Their worth has been suggested as an operating system for forecasting the European economies, not only in the realm of scientific discourse, but also in the domain of practical applicability, thereby laying the groundwork for further improvements. In addition, some of the conclusions obtained have contributed to an improved understanding of the workings of some of the major European economies.

§8.2 Main Contributions

The main contributions made in this work can be summarized as follows:

(a) A large macro data set is assembled from the currently existing IMF, OECD and ONS databanks through MIDAS, and then transformed and analysed for the European economies.

(b) Some forecasting records for the major European countries in the process of transition are examined, and the interactions among national economies are evaluated.

(c) Appropriate BVAR and BVAR-EC forecasting models for the UK, Germany, France and Italy are selected in both a single- and a multi-country context.

(d) Probability elicitation in Bayesian methodologies is conducted and programs using Bayesian networks are designed.

(e) Estimation and testing of unit roots and multicointegration in systems of equations are carried out.

(f) The best settings of the scalar prior hyperparameters are determined.

(g) Cointegration and BVAR techniques are combined.

(h) The rationality of forecasts in terms of their efficiency and unbiasedness properties is assessed, and the systematic comparison of alternative forecasting models is implemented.

(i) Forecast errors are decomposed into avoidable (systematic) and unavoidable (stochastic) components.

In terms of empirical practice, there are three main arguments in favour of Bayesian methods:

(a) BVAR models with a proper prior have a superiority in coping with nonstationarity in the data as the presence of unit roots cannot affect Bayesian asymptotics [cf. Phillips (1995c), p.93].

(b) BVAR models have inherent protection against overparameterization and can consistently outperform traditional methods by providing more precise forecasts in finite samples [cf. Koop & Potter (1997), p.10]. This highlights the importance of using BVAR in prediction.

(c) Long-run cointegration restrictions can reduce the sensitivity of BVAR forecasts across the values of prior hyperparameters, although imposing such restrictions may make little benefit for forecast accuracy, sometimes substantially so.

In conclusion, it may be said that the relatively simple BVAR models, which combine both prior and sample information in a flexible coherent manner, can lead to value-added in terms of forecast accuracy over classical methods, although as Artis *et al.* (1990b, p.16) point out:

> *the BVAR models must be 'tuned' correctly to achieve such results; inappropriate priors could yield less clear cut conclusions.*

In general, we concur with the argument that the BVAR forecasting models, on *a priori* grounds, can act as a highly effective standard of comparison for forecasts produced by more conventional methods.

§8.3 Problems

BVAR models have inherited the main advantages of UVARs in that they can generate unconditional forecasts without either designating which economic variables should be treated as underlying

determinants, or requiring the explicit theoretical restrictions. They can also avoid the overparameterization faced by any but very small UVAR models by incorporating the priors derived from statistical regularities. Unfortunately, it is these elaborated priors that put the BVAR approach under challenge.

The most commonly cited criticism of the BVAR models is that they are subjective, unscientific, or even capricious, relative to standard models, so that although they may capture the general tendency in the future, they cannot detect the significant cyclical turning points. Even if models of this class possess superior forecasting qualities, their usefulness would be in doubt if they could not enforce consistency or accommodate variation for policy or structural changes.

According to our records, no one model has ever given a perfect account of economic behaviour through time. There are a number of explanations for this. Given the complexity of 'real' economic systems, it is not surprising that forecasts are subject to errors. Some of the errors may be due to the stochastic nature of data and lack of knowledge of the parameters of the models. The parameters themselves may change over time in an inherently unpredictable way, so that the prior weight matrix may be both data- and time-dependent. All of these may lead to the deviation of forecasts from realized observations. However, one thing is certain. As long as the prior and cointegration restrictions are carefully chosen, and as long as the basic assumption that the same stochastic process will hold both during and after the estimation period is acceptable [Sargent (1984), pp.411-2], the gains from adopting BVAR and BVAR-EC models can be significant.

§8.4 Prospects for Future Developments

While the procedure developed here is likely to be useful, there are several directions in which it might be extended. The first of these

is to include an option to relax the restrictions of the standard Minnesota prior by specifying more general prior distributions that allow for dependencies between the VAR equations. In some cases, for example, economic variables are not well characterised as trending stochastically. Consequently, the standard Minnesota prior is not appropriate and an alternative prior is called for.

BVAR and BVAR-EC models are adaptive in nature. They represent only our best state of knowledge at a given time and use exclusively the previous information to adjust our ex ante forecasts mechanically. An intriguing avenue for further research would seem to be the development of the optimal forecast combination methodology using ancillary indicators in a Bayesian framework. In practice, much of the research to date demonstrates that forecast accuracy can be substantially improved through aggregating over multiple competing forecasts [Clemen (1989), p.559], a phenomenon supported by sound empirical results but which appears to be inconsistent with the encompassing approach. In theory, it should be possible to attain a single, catholic analytical model through the pooling of alternative information sets used by the component forecasts in the combination. This is certainly consistent with the encompassing approach but would appear to be either impossible or prohibitively costly.

In this respect, an eclectic but more comprehensive technique would be the combination of multiple forecasts with extraneous information developed by Fair & Shiller (1990), which entails treating forecasts themselves as pieces of information in a formal, general Bayesian framework. The motivation for building such a general Bayesian model by the effective use of additional information is, at its most basic level, the simple idea of the quest for 'rationality' in terms of both new information employed and increased forecast accuracy. This is, perhaps, a second best way forward but it is consistent with the literature of encompassing tests and of composite forecasts.

Another extension which may significantly improve forecast accuracy is the allowance of time-varying parameters. Doan *et al.* (1984) and Sims (1989) have done excellent work in constructing BVAR models

with time-varying parameters. Such an extension is likely to be most useful when the model appears to wander from the actual path. Furthermore, the extension of the methods to models with seasonal components, non-Gaussian and non-linear models for economic time series may also be sources of improving forecasts. Finally, 'impulse-response' policy simulations can be carried out in a multi-country context to determine how actual economies would react to specific shocks and, thereby, to investigate how a government would apply monetary and fiscal policies to pursue certain objectives in general environments.

Ultimately, it is anticipated that, with advances in econometric techniques, a fully formalised computable model could be established to deliver frequently optimal ex ante forecasts in the near future. In any case, it is essential that a macro model can interpret some macroeconomic episodes, fulfil consistency and parsimony criteria, and cope with both current and forward-looking information.

APPENDIX A

DOCUMENT EXAMINATION

The up-to-date document examination is carried out either through the Joint Academic NETwork (JANET) to the computer facilities of academic and other educational institutions world-wide, or through the NISS Gateway to all British library catalogues (OPACs). These facilities allow us access to data bases and other resources not available locally. Besides, we can also search via the Web (http://www.hull.ac.uk/lib/homepage.html) on Windows, or via the Social Science Citation Index (SSCI) in BIDS (Bath Information and Data Services) for articles and books (titles, citations, abstracts references) (published from 1960 forwards). This is normally done in two ways:

(1) search for journals in which the article has been published; and

(2) search for journals (or working papers) in which the article has been quoted (or cited).

The search results can then be sent directly to an email address and downloaded onto a floppy disk in plain text (*.txt) format.

APPENDIX B

MATHEMATICAL ELICITATION

This appendix provides the mathematical proof and elicitation of some important econometric propositions used in various chapters.

PROPOSITION: The first two moments of a stationary VAR(p) process, i.e. means, autocovariances and autocorrelations, can be written as:

$$\mu_Y = E[\underline{Y}_t] = \sum_{i=0}^{\infty} \underline{M}_i \underline{C};$$

$$\Gamma_y(s) = \underline{A}^s \Gamma_y(0), \quad \forall s \geq 0; \quad \text{and}$$

$$R_Y(s) = \underline{Q}^{-1} \Gamma_Y(s) \underline{Q}^{-1}, \quad \forall s \geq 0.$$

Here

$$\underline{M}_i = \sum_{j=1}^{\min(i,p)} \underline{M}_{i-j} \underline{A}_j, \quad \text{and } \underline{M}_0 = I_K;$$

$$\text{vec}\Gamma_y(0) = \left(I_{(Kp)^2} - \underline{A} \otimes \underline{A}\right)^{-1} \text{vec}\Sigma_u; \quad \text{and}$$

$$\underline{Q}^{-1} = \text{diag}\left[1/\sqrt{\gamma_{11}(0)}, \ldots, 1/\sqrt{\gamma_{KK}(0)}\right].$$

PROOF: Provided that the stationarity condition (2.5) is satisfied, then the means, autocovariances and autocorrelations of a stationary VAR(p) process may be derived by the following procedure.

Let

$$\underline{M}(L) \triangleq \sum_{i=0}^{\infty} \underline{M}_i L^i$$

be an operator such that

$$\underline{M}(L) \underline{A}_p(L) = I_K,$$

or

$$\underline{M}(L) = \underline{A}_p^{-1}(L). \tag{B.1}$$

Note also that Eq.(2.4) can be written alternatively as

$$\underline{Y}_t = \underline{A}_p^{-1}(L)\underline{C} + \underline{A}_p^{-1}(L)\underline{u}_t. \qquad (B.2)$$

Substituting (B.1) into (B.2) and rearranging yields

$$\begin{aligned}\underline{Y}_t &= \underline{M}(L)\underline{C} + \underline{M}(L)\underline{u}_t \\ &= \left(\sum_{i=0}^{\infty} \underline{M}_i\right)\underline{C} + \sum_{i=0}^{\infty} \underline{M}_i \underline{u}_{t-i}.\end{aligned}$$

This form of the process is called the infinite, or nonparsimonious, moving average (MA) representation, which expresses $\underline{Y}_t$ as an infinite distributed lag of the orthogonal innovations $\underline{u}_t$, and can be used to obtain the mean and autocovariances of $\underline{Y}_t$. That is,

$$\mu_Y = E[\underline{Y}_t] = \sum_{i=0}^{\infty} \underline{M}_i \underline{C} \qquad (B.3)$$

and

$$\begin{aligned}\Gamma_Y(s) &= E\left[(\underline{Y}_t - \mu_Y)(\underline{Y}_{t-s} - \mu_Y)'\right] \\ &= E\left[\left(\sum_{i=0}^{\infty} \underline{M}_i \underline{u}_{t-i}\right)\left(\sum_{i=0}^{\infty} \underline{M}_i \underline{u}_{t-s-i}\right)'\right] \\ &= E\left[\left(\sum_{i=0}^{s-1} \underline{M}_i \underline{u}_{t-i} + \sum_{i=0}^{\infty} \underline{M}_{i+s} \underline{u}_{t-s-i}\right)\left(\sum_{i=0}^{\infty} \underline{M}_i \underline{u}_{t-s-i}\right)'\right] \\ &= \sum_{i=0}^{\infty} \underline{M}_{i+s} \Sigma_u \underline{M}_i', \qquad s \geq 0,\end{aligned}$$

or (B.4)

$$\Gamma_Y(s) = \Gamma_Y(-s)', \qquad s < 0.$$

Here $\Sigma_u = E[\underline{u}_t \underline{u}_t']$ is again the white-noise variance-covariance matrix, and the M-weights can be drawn from the relationship (B.1), given A-weights, using the following recursions:

$$\begin{aligned}\underline{M}_0 &= I_K \\ \underline{M}_1 &= \underline{M}_0 \underline{A}_1 \\ \underline{M}_2 &= \underline{M}_1 \underline{A}_1 + \underline{M}_0 \underline{A}_2 \\ &\vdots \\ \underline{M}_i &= \sum_{j=1}^{\min(i,p)} \underline{M}_{i-j} \underline{A}_j. \qquad (B.5)\end{aligned}$$

For a stationary VAR(p) process, the M_i matrices will, in general,

approach zero as i tends to infinity. A more natural and attractive way of computing the autocovariances may be from the VAR coefficient matrices directly rather than from the MA coefficients, since (B.4) involves an infinite sum.

For expository purposes, any K-dimensional VAR(p) process can be stacked into a corresponding Kp-dimensional VAR(1) form after a suitable change of notation. The simple multivariate first-order VAR is fully general, i.e.,

$$\underline{y}_t = \underline{c} + \underline{\underline{A}}\underline{y}_{t-1} + \underline{u}_t, \tag{B.6}$$

where

$$\underline{y}_t \triangleq \underset{(Kp\times 1)}{\begin{pmatrix} \underline{y}_t \\ \underline{y}_{t-1} \\ \vdots \\ \underline{y}_{t-p-1} \end{pmatrix}}, \quad \underline{c} \triangleq \underset{(Kp\times 1)}{\begin{pmatrix} \underline{c} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}, \quad \underline{\underline{A}} \triangleq \underset{(Kp\times Kp)}{\begin{pmatrix} \underline{A}_1 & \underline{A}_2 & \cdots & \underline{A}_{p-1} & \underline{A}_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{pmatrix}}, \quad \underline{u}_t \triangleq \underset{(Kp\times 1)}{\begin{pmatrix} \underline{u}_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}}.$$

Since

$$\det\left[I_{Kp} - \underline{\underline{A}}\xi\right] = \det \underbrace{\begin{pmatrix} I_K - \underline{A}_1\xi & -\underline{A}_2\xi & -\underline{A}_3\xi & \cdots & -\underline{A}_{p-1}\xi & -\underline{A}_p\xi \\ -I_K\xi & I_K & 0 & \cdots & 0 & 0 \\ 0 & -I_K\xi & I_K & & 0 & 0 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & I_K & \\ 0 & 0 & & \cdots & -I_K\xi & I_K \end{pmatrix}}_{(Kp\times Kp)} \longleftarrow \text{expanding according to the cofactors of the first row}$$

$$= \det\left[I_K - \underline{A}_1\xi - \ldots - \underline{A}_p\xi^p\right] \neq 0, \qquad |\xi| \leq 1,$$

we know that $\underline{y}_t$ is also stationary and that the eigenvalues of $\underline{\underline{A}}$ have modulus less than 1. As such, the VAR(1) process may be written in mean-adjusted form as

$$\underline{y}_t - \underline{\mu}_y = \underline{\underline{A}}(\underline{y}_{t-1} - \underline{\mu}_y) + \underline{u}_t \tag{B.7}$$

with $\underline{\mu}_y \triangleq (\underline{\mu}_Y', \ldots, \underline{\mu}_Y')' = E[\underline{y}_t]$ and white noise covariance matrix $E[\underline{u}_t\underline{u}_t'] = \Sigma_u$. Postmultiplying (B.7) by $(\underline{y}_{t-s} - \underline{\mu}_y)'$ and taking

expectations gives

$$E[(\underline{y}_t - \underline{\mu}_y)(\underline{y}_{t-s} - \underline{\mu}_y)'] = \underline{\mathcal{A}}E[(\underline{y}_{t-1} - \underline{\mu}_y)(\underline{y}_{t-s} - \underline{\mu}_y)'] + E[\underline{u}_t(\underline{y}_{t-s} - \underline{\mu}_y)'].$$

Thus, for s=0

$$\Gamma_y(0) = \underline{\mathcal{A}}\Gamma_y(-1) + \Sigma_u = \underline{\mathcal{A}}\Gamma_y(1)' + \Sigma_u, \tag{B.8}$$

and for s>0

$$\Gamma_y(s) = \underline{\mathcal{A}}\Gamma_y(s-1). \tag{B.9}$$

These vector difference equations are often referred to as Yule-Walker equations with solution

$$\Gamma_y(s) = \underline{\mathcal{A}}^s\Gamma_y(0), \qquad s \geq 0, \tag{B.10}$$

where

$$\Gamma_y(s) = E\begin{bmatrix} \underline{Y}_t - \underline{\mu}_Y \\ \vdots \\ \underline{Y}_{t-p+1} - \underline{\mu}_Y \end{bmatrix}\left[(\underline{Y}_{t-s} - \underline{\mu}_Y)', \ldots, (\underline{Y}_{t-s-p+1} - \underline{\mu}_Y)'\right]$$

$$= \begin{bmatrix} \Gamma_Y(s) & \Gamma_Y(s+1) & \ldots & \Gamma_Y(s+p-1) \\ \Gamma_Y(s-1) & \Gamma_Y(s) & & \Gamma_Y(s+p-2) \\ \vdots & & & \vdots \\ \Gamma_Y(s-p+1) & \Gamma_Y(s-p+2) & \ldots & \Gamma_Y(s) \end{bmatrix}, \qquad s \geq 0.$$

(Kp×Kp)

If $\underline{\mathcal{A}}$ and Σ_u are known, the initial covariance matrix $\Gamma_y(0)$ can be determined as follows. For s=1 we have $\Gamma_y(1)=\underline{\mathcal{A}}\Gamma_y(0)$. Substituting $\Gamma_y(1)$ into (B.8) results in

$$\Gamma_y(0) = \underline{\mathcal{A}}\Gamma_y(0)\underline{\mathcal{A}}' + \Sigma_u$$

or

$$\text{vec}\Gamma_y(0) = \text{vec}\left(\underline{\mathcal{A}}\Gamma_y(0)\underline{\mathcal{A}}'\right) + \text{vec}\Sigma_u$$

$$= (\underline{\mathcal{A}} \otimes \underline{\mathcal{A}})\text{vec}\Gamma_y(0) + \text{vec}\Sigma_u.$$

where ⊗ denotes the matrix Kronecker direct product, while vec(·) signifies the column vectoring operator, which transforms the matrix

into a vector by stacking the columns one upon the other; hence,

$$\text{vec}\Gamma_y(0) = \left(I_{(Kp)^2} - \underline{A}\otimes\underline{A}\right)^{-1}\text{vec}\Sigma_u. \qquad (B.11)$$

The invertibility of $I_{(Kp)^2} - \underline{A}\otimes\underline{A}$ follows from the stationarity of $\underline{y}_t$ because the eigenvalues of $\underline{A}\otimes\underline{A}$, being the products of the eigenvalues of $\underline{A}$, must have modulus less than 1. Accordingly, $\det\left[I_{(Kp)^2} - \underline{A}\otimes\underline{A}\right] \neq 0$.

Furthermore, the unit-dependent autocovariance matrices may be standardised to yield the corresponding unit-free autocorrelation matrices

$$R_Y(s) = \underline{Q}^{-1}\Gamma_Y(s)\underline{Q}^{-1}, \qquad (B.12)$$

which are generally more convenient to work with. Here $\underline{Q}$ is a diagonal matrix with the square roots of the diagonal elements of $\Gamma_Y(0)$ on its diagonal. Denoting the ij-th elements of $\Gamma_Y(s)$ (or the covariance between $Y_{i,t}$ and $Y_{j,t-s}$) as $\gamma_{ij}(s)$ and the diagonal elements of $\Gamma_Y(0)$ (or the variances of $Y_{1t},\ldots,Y_{Kt}$) as $\gamma_{ij}(0)$, we have

$$\underline{Q}^{-1} = \text{diag}\left[1/\sqrt{\gamma_{11}(0)},\ldots,1/\sqrt{\gamma_{KK}(0)}\right].$$

The ij-th element of $R_Y(s)$ signifies the auto- or cross-correlation between a pair of elements, Y_i and Y_j, in the vector $\underline{Y}$ at lag s, i.e.,

$$\rho_{ij}(s) = \gamma_{ij}(s)/\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)},$$

and therefore the temporal (dual) symmetric crosscorrelation function (CCF) between Y_i and Y_j, say, is $\rho_{ij}(s)\left(=\rho_{ij}(-s)\right)$, $\forall s\geq 0$.

Q.E.D.

In practice, the easiest way to remove $\underline{C}$ is to convert $\underline{Y}_t$, observed for t=1, ..., T periods, into deviations from its sample mean:

$$\underline{A}_p(L)\left(\underline{Y}_t - \bar{\underline{Y}}\right) = \left(\underline{U}_t - \bar{\underline{U}}\right),$$

given that the process is stationary. This can be done simply by stacking the observations into a KT-vector $\underline{Y}=(Y_{1,1},\ldots,Y_{1,T},\ldots,Y_{K,1},\ldots,Y_{K,T})'$:

$$(\underline{A}_p(L)\otimes I_T)\underline{Y} = \underline{C}\otimes J_T + \underline{U},$$

where J_T is a T-rack identity matrix, and J_T a T-vector of ones. Define

$$\underline{Q}_T \equiv I_T - \frac{1}{T} J_T J_T',$$

which is symmetric and idempotent. Premultiplying by $I_K\otimes\underline{Q}_T$ is equivalent to scaling the equations in terms of deviations about the means:

$$(\underline{A}_p(L)\otimes\underline{Q}_T)\underline{Y} = (I_K\otimes\underline{Q}_T)\underline{U}.$$

Here $(I_K\otimes\underline{Q}_T)(\underline{A}_p(L)\otimes I_T) = \underline{A}_p(L)\otimes\underline{Q}_T$ and

$(I_K\otimes\underline{Q}_T)(\underline{C}\otimes J_T) = \underline{C}\otimes(\underline{Q}_T J_T) = 0$; as $\underline{Q}_T J_T = 0$.

PROPOSITION: The marginal or predictive pdf for $\tilde{\underline{Y}}$, $f(\tilde{\underline{Y}}|\underline{Y})$, in (3.22) can be written as:

$$f(\tilde{\underline{Y}}|\underline{Y}) \propto |[\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\overline{\underline{\beta}}]'\underline{W}[\tilde{\underline{Y}}-(I_K\otimes\tilde{\underline{X}})\overline{\underline{\beta}}]|^{-\gamma/2}, \qquad (B.13)$$

where

$$\underline{W} \triangleq (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'](\Sigma_u^{-1/2}\otimes I_h);$$
$$\underline{N} \triangleq \overline{\underline{V}}_\beta^{-1}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}}); \text{ and}$$
$$\gamma = T+h-(1+Kp).$$

PROOF: To integrate the unknown parameter vector $\underline{\beta}$ out of the predictive pdf for $\tilde{\underline{Y}}$, $f(\tilde{\underline{Y}}|\underline{Y})$, in (3.22):

$$f(\tilde{\underline{Y}}|\underline{Y}) \propto \int\exp\Big\{-\frac{1}{2}[(\overline{\underline{V}}_\beta^{-1/2}(\underline{\beta}-\overline{\underline{\beta}}))'(\overline{\underline{V}}_\beta^{-1/2}(\underline{\beta}-\overline{\underline{\beta}}))$$
$$+ ((\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{\beta})'((\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{\beta})]\Big\}d\underline{\beta}, \qquad (B.14)$$

we define

$$m \triangleq \begin{bmatrix} \bar{V}_\beta^{-1/2}\bar{\beta} \\ (\Sigma_u^{-1/2}\otimes I_h)\tilde{Y} \end{bmatrix} \text{ and } M \triangleq \begin{bmatrix} \bar{V}_\beta^{-1/2} \\ \Sigma_u^{-1/2}\otimes\tilde{X} \end{bmatrix},$$

and complete the square on β for the integrand. This gives

$$\begin{aligned} &\exp\left\{-\tfrac{1}{2}(m-M\beta)'(m-M\beta)\right\} \\ &= \exp\left\{-\tfrac{1}{2}(m'm+\beta'M'M\beta-\beta'M'm-m'M\beta)\right\} \\ &= \exp\left\{-\tfrac{1}{2}(m'm+(\beta-(M'M)^{-1}M'm)'M'M \right. \\ &\qquad \left.(\beta-(M'M)^{-1}M'm)-m'M(M'M)^{-1}M'm)\right\}. \end{aligned}$$

On substituting this expression in (B.14), properties of the Wishart pdf can be utilized to integrate with respect to the K(1+Kp) elements of β, given Σ_u, which yields the predictive pdf for $\tilde{Y}$, i.e.

$$\begin{aligned} f(\tilde{Y}|Y) \propto\ & \frac{1}{|m'm-m'M(M'M)^{-1}M'm|^{\gamma/2}} \\ & \times\int\left\{\exp\left\{-\tfrac{1}{2}\left[m'm-m'M(M'M)^{-1}M'm+\right.\right.\right. \\ & \qquad \left.\left.\left.(\beta-(M'M)^{-1}M'm)'M'M(\beta-(M'M)^{-1}M'm)\right]\right\}\Big/ \\ & \qquad |m'm-m'M(M'M)^{-1}M'm|^{-\gamma/2}\right\}d\beta, \end{aligned}$$

or

$$f(\tilde{Y}|Y) \propto |m'm-m'M(M'M)^{-1}M'm|^{-\gamma/2};\ \text{for } \gamma = T+h-(1+Kp),$$

since the accumulated new data points are allowed to assist in selecting the most appropriate forecasting model and the above integral is just equal to the normalizing constant of the Wishart pdf which does not depend on the parameter vector β.[1]

To put (B.14) in a more convenient form, we define

$$N \triangleq M'M = \bar{V}_\beta^{-1} + (\Sigma_u^{-1/2}\otimes\tilde{X})'(\Sigma_u^{-1/2}\otimes\tilde{X}),$$

[1]For details, see Zellner (1971), Chapter VIII.

and complete the square on $\tilde{\underline{Y}}$ as follows:

$$
\begin{aligned}
&\underline{m}'\underline{m}-\underline{m}'\underline{M}(\underline{M}'\underline{M})^{-1}\underline{M}'\underline{m} \\
&= \bar{\underline{\beta}}'\bar{\underline{V}}_\beta^{-1}\bar{\underline{\beta}}+\tilde{\underline{Y}}'(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}} - \\
&\quad [\bar{\underline{\beta}}'\bar{\underline{V}}_\beta^{-1}+\tilde{\underline{Y}}'(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})]\underline{N}^{-1}[\bar{\underline{V}}_\beta^{-1}\bar{\underline{\beta}}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}}] \\
&= \bar{\underline{\beta}}'(\bar{\underline{V}}_\beta^{-1}-\bar{\underline{V}}_\beta^{-1}\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1})\bar{\underline{\beta}} + \\
&\quad \tilde{\underline{Y}}'(\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'](\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}} - \\
&\quad \bar{\underline{\beta}}'\bar{\underline{V}}_\beta^{-1}\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes I_h)\tilde{\underline{Y}} - \tilde{\underline{Y}}'(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1}\bar{\underline{\beta}}.
\end{aligned}
\tag{B.15}
$$

Let

$$
\underline{W} \triangleq (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'](\Sigma_u^{-1/2}\otimes I_h),
$$

the quantity in the third line of (B.15) can be expressed as

$$
\begin{aligned}
&\bar{\underline{\beta}}'[\bar{\underline{V}}_\beta^{-1}-\bar{\underline{V}}_\beta^{-1}\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1}-\bar{\underline{V}}_\beta^{-1}\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes I_h)\underline{W}^{-1}(\Sigma_u^{-1/2}\otimes I_h)' \\
&(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1}]\bar{\underline{\beta}} + [\tilde{\underline{Y}}-\underline{W}^{-1}(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1}\bar{\underline{\beta}}]'\underline{W} \\
&[\tilde{\underline{Y}}-\underline{W}^{-1}(\Sigma_u^{-1/2}\otimes I_h)'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}\bar{\underline{V}}_\beta^{-1}\bar{\underline{\beta}}].
\end{aligned}
\tag{B.16}
$$

Further, we have

$$
\underline{W}^{-1} = (\Sigma_u^{-1/2}\otimes I_h)^{-1}[I_{Kh}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\bar{\underline{V}}_\beta(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'][(\Sigma_u^{-1/2}\otimes I_h)']^{-1}.^{2}
$$

---

[2] This result can be verified by direct multiplication:

$$
\begin{aligned}
\underline{W}\underline{W}^{-1} &= (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'] \\
&\qquad [I_{Kh}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\bar{\underline{V}}_\beta(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'][(\Sigma_u^{-1/2}\otimes I_h)']^{-1} \\
&= (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})(\underline{N}^{-1}-\bar{\underline{V}}_\beta+\underline{N}^{-1}(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})' \\
&\qquad (\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\bar{\underline{V}}_\beta)(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'][(\Sigma_u^{-1/2}\otimes I_h)']^{-1} \\
&= (\Sigma_u^{-1/2}\otimes I_h)'[I_{Kh}-(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})\underline{N}^{-1}(\bar{\underline{V}}_\beta^{-1}-\underline{N}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})) \\
&\qquad \bar{\underline{V}}_\beta(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'][(\Sigma_u^{-1/2}\otimes I_h)']^{-1} \\
&= (\Sigma_u^{-1/2}\otimes I_h)'[(\Sigma_u^{-1/2}\otimes I_h)']^{-1} \\
&= I_{Kh}, \quad \text{since } \bar{\underline{V}}_\beta^{-1}-\underline{N}+(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})'(\Sigma_u^{-1/2}\otimes\tilde{\underline{X}})=0.
\end{aligned}
$$

Also

$$\begin{aligned}
&\underline{W}^{-1}\left(\Sigma_u^{-1/2}\otimes I_h\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\mathcal{N}}^{-1}\\
&= \left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left[I_{Kh}+\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\right]\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\mathcal{N}}^{-1}\\
&= \left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\left[I_{Kh}+\underline{\bar{V}}_\beta\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\right]\underline{\mathcal{N}}^{-1}\\
&= \left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta\left[\underline{\bar{V}}_\beta^{-1}+\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\right]\underline{\mathcal{N}}^{-1}\\
&= \left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta, \quad \text{since } \underline{\mathcal{N}} \triangleq \underline{\bar{V}}_\beta^{-1}+\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right).
\end{aligned} \tag{B.17}$$

Finally

$$\begin{aligned}
&\underline{\bar{V}}_\beta^{-1}-\underline{\bar{V}}_\beta^{-1}\underline{\mathcal{N}}^{-1}\underline{\bar{V}}_\beta^{-1}-\underline{\bar{V}}_\beta^{-1}\underline{\mathcal{N}}^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes I_h\right)\underline{W}^{-1}\\
&\qquad\left(\Sigma_u^{-1/2}\otimes I_h\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\mathcal{N}}^{-1}\underline{\bar{V}}_\beta^{-1}\\
&= \underline{\bar{V}}_\beta^{-1}-\underline{\bar{V}}_\beta^{-1}\left[\underline{\mathcal{N}}^{-1}+\underline{\mathcal{N}}^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta\right]\underline{\bar{V}}_\beta^{-1}\\
&= \underline{\bar{V}}_\beta^{-1}-\underline{\bar{V}}_\beta^{-1}\underline{\mathcal{N}}^{-1}\left[\underline{\bar{V}}_\beta^{-1}+\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\right]\\
&= \underline{\bar{V}}_\beta^{-1}-\underline{\bar{V}}_\beta^{-1} = 0, \quad \text{since } \underline{\mathcal{N}} \triangleq \underline{\bar{V}}_\beta^{-1}+\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)'\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right).
\end{aligned} \tag{B.18}$$

Substitution from (B.17) and (B.18) in (B.16) leads to

$$\begin{aligned}
&\left[\underline{\tilde{Y}}-\left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta\underline{\bar{V}}_\beta^{-1}\underline{\bar{\beta}}\right]'\underline{W}\\
&\qquad\left[\underline{\tilde{Y}}-\left(\Sigma_u^{-1/2}\otimes I_h\right)^{-1}\left(\Sigma_u^{-1/2}\otimes\underline{\tilde{X}}\right)\underline{\bar{V}}_\beta\underline{\bar{V}}_\beta^{-1}\underline{\bar{\beta}}\right]\\
&= \left[\underline{\tilde{Y}}-\left(I_K\otimes\underline{\tilde{X}}\right)\underline{\bar{\beta}}\right]'\underline{W}\left[\underline{\tilde{Y}}-\left(I_K\otimes\underline{\tilde{X}}\right)\underline{\bar{\beta}}\right].
\end{aligned}$$

As it stands, the predictive pdf (3.23) can be readily obtained

$$f\left(\underline{\tilde{Y}}|\underline{Y}\right) \propto \left|\left[\underline{\tilde{Y}}-\left(I_K\otimes\underline{\tilde{X}}\right)\underline{\bar{\beta}}\right]'\underline{W}\left[\underline{\tilde{Y}}-\left(I_K\otimes\underline{\tilde{X}}\right)\underline{\bar{\beta}}\right]\right|^{-\gamma/2}. \tag{B.19}$$

Q.E.D.

APPENDIX C

DATABASE ASSEMBLY

This appendix only presents the macro data used in this work. The data sets are assembled from the currently existing OECD Main Economic Indicators, IMF International Financial Statistics (IFS) and ONS Macro-Economic Time Series databanks in Manchester Computing Centre (MCC) through MIDAS. The complete macro data for the four leading (and other) European economies are stored in ASCII files ending with suffix *.ASC in an accompanying data diskette.

UNITED KINGDOM

(I) IMF IFS DATABASE

```
* MONEY (Billions of Pounds: End of Period)
* periodicity: QUARTERLY  seasonally: UNADJUSTED  starts: 1960Q1
* stops: 1994Q4  no. of obs: 137  last updated: 1994Q1
* one line of data represents two years

  10.894  10.808  11.551  11.246    11.202  11.166  11.045  11.473
  10.784  11.110  11.254  10.852    11.322  11.633  11.789  12.425
  12.126  12.257  12.517  12.824    12.486  12.661  12.809  13.318
  13.175  13.114  13.246  13.311    13.190  13.404  13.957  14.327
  13.932  14.180  13.738  14.906    14.151  13.895  14.105  14.953
  14.436  15.021  15.327  16.350    16.445  16.683  17.326  18.816
  19.048  19.903  20.245  21.478    20.928  22.357  21.860  22.574
  21.673  22.357  22.931  25.011    25.004  26.983  28.463  29.666
  30.207  31.037  32.625  33.034    33.202  34.635  37.416  39.917
  40.908  41.665  43.866  46.435    46.325  46.924  48.748  50.664
  49.176  50.082  50.174  52.680    53.354  55.217  56.152  58.704
  57.694  63.233  64.827  68.993    70.790  72.643  73.401  72.056
  79.465  82.996  85.433  81.537    91.662  95.696 100.196  96.173
 110.003 116.808 125.235 117.539   138.501 145.111 149.012 154.118
 159.933 166.270 170.086 170.673   175.037 185.919 191.741 195.309
 198.868 205.509 207.334 214.940   217.140 225.180 227.180 229.220
 232.390 235.590 237.950 238.950   242.350 243.240 245.210 252.340
 254.920   0.000   0.000   0.000

* QUASI-MONEY (Billions of Pounds: End of Period)
* periodicity: QUARTERLY  seasonally: UNADJUSTED  starts: 1960Q1
```

* stops: 1994Q4 no. of obs: 125 last updated: 1994Q1
* one line of data represents two years

0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	5.421	5.670	5.726	5.981
5.846	6.146	6.249	6.548	6.476	6.945	7.058	7.453
7.428	7.689	7.812	8.139	7.981	8.423	8.639	9.375
9.332	9.885	10.494	10.436	10.591	10.754	10.859	11.120
10.865	11.454	11.776	12.181	12.080	12.596	12.748	13.526
14.515	16.182	17.616	19.535	21.150	22.114	25.946	29.075
30.219	30.358	32.239	33.251	32.768	32.184	32.982	33.144
32.618	34.582	36.270	37.038	35.896	37.807	36.771	37.314
38.369	41.182	40.715	42.142	41.420	45.454	45.765	48.760
49.985	55.545	59.152	64.104	65.235	72.445	78.081	86.493
87.186	91.827	93.505	96.341	99.792	103.255	104.747	109.694
108.910	109.958	112.854	121.289	119.473	122.129	126.472	130.206
136.295	145.319	153.951	159.873	167.971	174.427	182.675	189.073
194.875	204.037	219.808	231.932	245.310	258.140	268.960	286.060
299.990	309.400	320.500	318.530	319.930	325.740	326.490	302.740
310.500	312.310	319.130	326.660	328.250	334.350	339.940	349.600
356.970	0.000	0.000	0.000				

* MONEY PLUS QUASI-MONEY (Billions of Pounds: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 135 last updated: 1996Q3
* one line of data represents two years

0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	16.743	17.302	17.514	18.406
17.972	18.403	18.765	19.371	18.962	19.606	19.866	20.771
20.603	20.803	21.059	21.450	21.171	21.828	22.596	23.702
23.264	24.064	24.233	25.342	24.742	24.648	24.964	26.073
25.301	26.475	27.103	28.531	28.525	29.278	30.074	32.342
33.564	36.085	37.861	41.013	42.078	44.471	47.806	51.649
51.893	52.716	55.169	58.263	57.772	59.168	61.445	62.810
62.826	65.619	68.896	70.072	69.099	72.442	74.187	77.231
79.277	82.847	84.581	88.577	87.745	92.378	94.513	99.424
99.161	105.627	109.325	116.784	118.589	127.662	134.233	145.197
144.881	155.060	158.332	165.333	170.582	175.898	178.148	181.750
188.375	192.954	198.286	202.825	211.135	217.824	226.668	226.379
246.298	262.127	279.186	277.412	306.472	319.538	331.687	343.191
354.808	370.307	389.894	402.605	420.347	444.059	460.701	481.369
498.858	514.909	527.834	533.470	537.070	550.920	553.670	531.960
542.890	547.900	573.790	582.810	585.730	590.910	596.190	610.990
626.310	633.840	637.660	646.470	668.930	683.990	700.800	729.630
759.890	783.330	796.210	0.000				

* M0 (Billions of Pounds: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 110 last updated: 1996Q3
* one line of data represents two years

0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	3.819	3.818	3.993

3.834	3.943	4.045	4.206	4.289	4.322	4.307	4.425
4.316	4.583	4.638	5.053	4.888	5.096	5.187	5.593
5.420	5.575	5.886	6.444	6.341	6.535	6.689	7.228
6.956	7.279	7.534	7.988	7.700	8.139	8.256	9.122
8.898	9.225	9.528	10.362	9.989	10.352	10.725	11.620
10.994	11.295	11.523	12.243	11.755	11.931	11.889	12.555
11.729	11.960	12.155	12.948	12.431	12.678	12.830	13.849
13.030	13.339	13.575	14.615	13.738	14.023	14.090	15.161
14.440	14.509	14.755	16.098	14.809	15.083	15.533	16.633
15.751	16.247	16.793	18.040	16.815	17.089	17.504	19.006
17.600	18.194	18.330	19.490	18.180	18.590	18.750	20.090
18.410	18.850	19.230	20.580	19.310	19.840	20.270	21.730
20.560	21.160	21.610	23.320	21.880	22.270	22.800	24.540
23.120	23.930	24.440	0.000				

* CONSUMER PRICE INDEX (CPI): PERIOD AVERAGES (1990=100)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 147 last updated: 1996Q3
* one line of data represents two years

9.8	9.8	9.8	9.9	10.0	10.1	10.2	10.2
10.4	10.7	10.7	10.7	10.9	10.9	10.8	10.9
11.0	11.2	11.3	11.4	11.5	11.7	11.8	11.9
11.9	12.3	12.3	12.4	12.4	12.5	12.5	12.6
12.8	13.1	13.2	13.3	13.6	13.8	13.8	13.9
14.3	14.6	14.8	15.1	15.5	16.0	16.3	16.5
16.7	17.0	17.3	17.7	18.1	18.6	18.9	19.6
20.4	21.6	22.1	23.1	24.5	26.8	28.0	29.0
30.0	31.1	31.8	33.3	35.0	36.5	37.1	37.6
38.3	39.3	40.0	40.7	42.0	43.5	46.4	47.7
50.0	52.9	54.0	55.0	56.3	59.1	60.1	61.6
62.6	64.6	64.9	65.4	65.7	67.1	67.9	68.7
69.1	70.5	71.1	72.0	72.9	75.4	75.6	76.0
76.5	77.6	77.6	78.6	79.6	80.8	81.0	81.8
82.2	84.2	85.4	87.2	88.6	91.1	92.0	93.8
95.5	99.9	101.6	103.1	103.7	105.9	106.4	107.4
108.0	110.3	110.2	110.7	109.9	111.7	112.0	112.4
112.6	114.6	114.6	115.4	116.4	118.5	118.8	119.0
119.6	121.2	121.4	0.0				

* TREASURY BILL RATE (Percent Per Annum)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 147 last updated: 1996Q3
* one line of data represents two years

4.40	4.71	5.56	4.85	4.35	4.45	6.14	5.57
5.21	4.00	3.79	3.71	3.51	3.69	3.72	3.72
3.98	4.36	4.62	5.47	6.51	6.12	5.56	5.45
5.56	5.65	6.58	6.62	5.99	5.30	5.44	6.57
7.39	7.15	6.95	6.65	7.19	7.83	7.81	7.71
7.54	6.89	6.83	6.82	6.74	5.67	5.39	4.52
4.36	4.59	5.96	7.15	8.13	7.36	10.24	11.62
11.99	11.36	11.18	10.96	10.01	9.39	10.17	11.13
9.04	10.19	11.26	13.99	11.20	7.72	6.55	5.26
5.90	8.11	6.34	11.00	11.94	11.75	13.34	14.87
16.04	16.02	14.60	13.99	12.16	11.59	13.55	14.82
13.37	12.57	10.61	9.33	10.58	9.64	9.29	8.85

8.72	8.69	10.40	9.39	12.06	11.92	11.14	11.10
11.67	9.67	9.48	10.64	10.05	8.81	9.44	8.69
8.49	7.92	10.75	11.96	12.42	12.87	12.45	14.47
14.47	14.47	14.30	13.08	12.83	10.89	10.13	9.98
9.94	9.60	9.41	6.78	5.60	5.27	5.12	4.99
4.82	4.86	5.27	5.65	6.06	6.29	6.58	6.41
5.95	5.73	5.52	0.00				

* GROSS DOM. PROD. 1990 PRICES (Billions of Pounds)
* periodicity: QUARTERLY seasonally: ADJUSTED (at Annual Rates and
* May Not Average to Yearly Data) starts: 1960Q1 stops: 1996Q4
* no. of obs: 42 last updated: 1996Q2
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
476.23	480.91	484.88	490.52	495.73	501.22	511.98	516.98
522.71	539.19	533.18	537.85	540.43	542.73	545.62	547.52
554.64	557.65	551.74	546.73	543.12	539.87	538.66	539.58
535.83	536.10	538.68	539.18	543.88	545.31	550.79	554.36
560.39	567.34	573.07	577.73	580.28	582.42	584.93	587.83
591.36	594.01	598.44	0.00				

* GROSS DOM. PROD. 1985 PRICES (Billions of Pounds)
* periodicity: QUARTERLY seasonally: UNADJUSTED (at Annual Rates
* and May Not Average to Yearly Data) starts: 1960Q1 stops: 1988Q4
* no. of obs: 65 last updated: 1988Q1
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
64.97	69.65	69.03	73.29	70.62	72.16	72.98	76.27
68.56	72.10	74.49	77.50	69.42	70.45	71.34	75.52
72.93	72.38	75.47	79.56	73.66	74.21	76.55	81.62
76.72	77.36	79.70	83.16	77.45	79.54	81.79	84.05
78.64	76.72	79.62	81.65	79.04	77.06	79.52	82.74
79.67	77.78	80.50	84.15	82.40	79.41	84.08	86.98
83.94	81.73	85.38	90.09	86.40	85.81	89.11	92.39
88.81	87.45	90.74	96.94	91.38	89.53	95.60	100.39
93.69	0.00	0.00	0.00				

(II) OECD MEI DATABASE

```
* GROSS DOMESTIC PRODUCT: CONSTANT PRICES OF 1990
* (Billions of Pounds, Market Prices, Annual Rates)
* periodicity: QUARTERLY  seasonally: ADJUSTED  starts: 60Q1
* stops: 96Q4  no. of obs: 148  last updated: 96Q4
* one line of data represents two years

 263.160 258.807 262.987 266.208    271.430 271.938 272.441 270.199
 270.424 276.055 275.500 275.562    276.547 287.211 286.787 293.250
 295.995 300.735 299.206 306.984    305.444 305.198 307.654 311.567
 311.153 313.416 313.777 317.328    322.629 321.911 324.656 321.241
 334.828 329.784 337.797 341.946    335.037 338.902 342.637 345.628
 341.123 348.964 350.609 354.652    348.262 356.863 364.311 364.567
 360.665 367.113 365.913 375.619    397.890 395.837 396.701 391.479
 386.686 391.086 397.298 392.715    390.667 389.709 384.088 392.485
 401.991 398.162 402.834 410.680    403.725 403.416 406.925 416.934
 417.552 420.810 422.685 428.536    419.883 442.940 431.500 434.685
 429.961 422.560 427.326 420.349    421.376 419.899 419.967 421.145
 424.094 426.027 424.938 429.966    439.640 439.886 443.395 447.496
 450.492 448.099 447.669 453.352    461.743 466.997 467.777 469.909
 477.210 481.699 487.906 493.499    497.203 502.477 513.314 518.337
 521.884 524.621 529.845 536.143    537.582 538.338 542.528 547.848
 550.388 553.003 545.555 540.180    540.702 538.066 537.508 538.448
 533.574 536.367 538.442 539.660    543.600 546.610 552.300 555.840
 561.200 567.340 573.070 577.730    579.600 582.770 584.600 588.300
 591.290 593.930 596.560 603.380

* IMPLICIT PRICE LEVEL: 1990=100
* periodicity: QUARTERLY  seasonally: ADJUSTED  starts: 60Q1
* stops: 96Q4  no. of obs: 148  last updated: 96Q4
* one line of data represents two years

   9.6     9.8     9.8     9.9      10.0     9.9    10.3    10.2
  10.4    10.5    10.5    10.6      10.6    10.7    10.8    10.9
  10.8    11.1    11.2    11.3      11.5    11.6    11.8    11.8
  12.0    12.1    12.3    12.4      12.4    12.6    12.6    12.7
  12.8    13.1    13.2    13.4      13.6    13.7    13.9    14.0
  14.4    14.6    15.0    15.3      15.7    16.0    16.4    16.7
  17.0    17.3    17.6    18.1      18.4    18.3    18.8    19.5
  19.8    21.0    22.0    23.4      25.1    27.0    28.3    29.3
  30.2    31.0    31.9    33.1      34.2    35.6    36.5    37.3
  38.6    39.6    40.4    41.4      42.8    44.0    47.3    49.0
  51.4    53.8    55.9    57.9      59.3    60.3    61.3    62.9
  64.1    64.9    66.0    67.3      68.1    68.2    69.3    70.5
  70.7    71.9    72.5    73.6      74.7    75.9    76.9    78.0
  78.3    78.4    79.0    80.1      81.0    82.4    83.5    84.8
  85.7    87.6    89.3    91.0      92.7    94.0    95.0    96.4
  98.5    99.0   101.9   102.9     103.9   106.3   107.2   108.6
 109.6   111.3   111.7   111.9     113.5   114.2   115.4   116.0
 116.5   116.7   117.7   117.8     118.8   119.8   120.1   120.9
 122.3   123.0   124.3   124.9

* MONEY SUPPLY M1 (Billions of Pounds)
* periodicity: QUARTERLY  seasonally: UNADJUSTED  starts: 1960Q1
* stops: 1993Q4  no. of obs: 75  last updated: 1990Q3
```

* one line of data represents two years

0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5.968	7.213	8.395	9.183	10.673	10.891	12.843	13.874
16.721	17.000	17.804	18.180	18.710	18.465	19.133	19.543
19.612	20.371	21.182	22.423	22.385	23.108	23.820	24.247
24.612	26.051	26.047	26.931	27.741	28.619	29.270	30.758
31.714	33.036	35.177	35.901	37.255	39.314	43.133	47.211
53.936	58.192	61.844	65.805	71.672	73.419	77.060	82.238
86.100	90.684	95.006	102.955	112.767	114.522	114.927	118.697
122.277	129.989	137.682	149.883	157.659	171.342	180.908	190.749
196.186	206.873	221.554	233.364	250.335	267.044	309.081	324.293
331.887	342.099	345.974	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

* IMPORTS C.I.F. (millions of pounds, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 148 last updated: 1996Q4
* one line of data represents two years

374	388	390	400	392	379	369	375
378	384	393	388	390	409	423	439
470	470	473	485	456	482	488	491
504	492	510	478	525	524	526	570
644	644	669	676	674	699	695	703
720	770	751	796	791	830	809	836
872	889	882	1049	1132	1225	1367	1517
1735	1968	1991	2019	1980	1875	2062	2099
2180	2576	2696	2909	2959	3145	3017	2951
3115	3333	3357	3373	3473	4000	3945	4223
4451	4351	3981	3807	3617	3959	4722	4758
4653	4892	4695	4753	5299	5244	5505	5786
5940	6437	6712	7200	7521	7182	6774	6785
6914	6829	7271	7756	7273	7606	8329	8289
8042	8635	9331	9412	9655	9966	10438	10226
10718	10802	10210	10407	9724	9822	9998	10062
10164	10404	10351	11047	11089	11195	11477	11628
12076	12282	12447	13158	13297	13964	14434	14844
15418	15437	15281	15176				

* EXPORTS F.O.B. (millions of pounds, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 148 last updated: 1996Q4
* one line of data represents two years

322	314	310	317	330	327	333	328
328	342	339	345	350	360	369	377
376	384	375	388	394	401	415	424
428	417	442	465	463	448	436	396
519	510	552	563	562	600	640	645
668	660	639	732	683	678	664	797
767	795	709	930	922	969	1039	1099

1205	1345	1437	1450	1569	1551	1630	1786
1873	2054	2139	2359	2496	2656	2819	2693
2791	2953	2976	3074	2782	3570	3500	3693
3986	3979	3912	3911	3929	4057	4358	4556
4502	4622	4590	4806	4881	4893	5088	5366
5569	5632	5864	6431	6716	6753	6276	6366
6086	5938	5887	6449	6546	6439	6784	6922
6316	6798	7098	6858	7302	7550	7986	8372
8511	8787	8644	8722	8431	8743	8904	8871
8817	9045	8917	9320	9865	9735	9963	10114
10679	11092	11500	11792	12579	12601	13077	13400
13826	13799	14069	13831				

(III) CSO MTS DATABANK (SOURCE: BANK OF ENGLAND)

* M0 - THE WIDE MONETARY BASE: AMOUNT OUTSTANDING (AVER: £ million)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 111 last updated: 1996Q4
* one line of data represents two years

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	3785	3803	3917
3882	3915	4042	4097	4348	4299	4314	4289
4376	4564	4655	4845	4976	5088	5200	5339
5525	5586	5882	6157	6369	6541	6686	6872
7081	7272	7511	7648	7826	8127	8280	8651
8949	9276	9571	9836	10136	10420	10782	11013
11151	11376	11566	11649	11946	12018	11893	11925
11911	11993	12144	12302	12550	12702	12850	13038
13230	13401	13591	13746	13959	14093	14083	14266
14461	14570	14751	15022	15103	15203	15463	15666
15978	16279	16761	16866	16969	17168	17468	17824
18009	18253	18288	18300	18455	18649	18722	18857
18846	18943	19144	19394	19695	19850	20182	20561
20819	21185	21609	21968	22304	22355	22752	23206
23540	24012	24356	24824				

(I) IMF IFS DATABASE

* MONEY SUPPLY M1 (billions of Deutsche Mark)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 147 last updated: 1996Q3
* one line of data represents two years

43.2	44.6	45.3	47.4	46.3	48.6	50.4	54.4
51.6	53.9	55.2	58.0	55.0	57.5	59.0	62.2
59.6	62.8	63.8	67.6	65.6	69.1	69.4	72.8
69.4	72.3	71.9	74.2	70.8	73.5	75.0	80.5
78.3	81.8	83.2	90.6	84.8	89.9	91.2	95.4
90.7	95.5	96.5	103.7	99.4	106.9	109.5	116.9
114.4	123.0	124.8	133.3	128.9	128.8	125.1	135.7
128.7	136.0	137.4	150.3	144.9	155.7	160.1	171.6
162.1	173.7	171.8	177.3	173.3	183.1	186.5	198.6
197.4	208.4	211.0	227.5	218.0	224.6	222.0	234.2
221.6	228.8	229.2	243.4	224.9	233.2	226.5	239.6
228.8	240.2	238.8	256.7	253.0	266.7	263.2	278.2
263.1	271.5	270.6	294.8	272.5	281.2	284.8	314.5
299.2	313.4	315.7	340.2	324.4	344.7	344.5	365.7
356.9	379.4	376.8	408.3	390.9	398.5	394.5	431.6
400.5	463.8	480.4	551.4	508.9	518.8	527.2	575.0
535.3	555.7	565.1	641.0	590.7	610.4	622.6	697.6
650.1	676.8	675.9	732.0	675.6	697.4	704.7	783.7
751.5	772.6	784.7	0.0				

* QUASI-MONEY (billions of Deutsche Mark)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 147 last updated: 1996Q3
* one line of data represents two years

55.1	56.1	57.7	60.4	63.3	64.0	64.8	67.3
70.3	71.5	72.7	76.6	80.1	81.4	83.6	88.6
92.2	94.0	96.5	102.5	106.9	109.8	111.7	119.8
123.6	126.7	130.9	139.7	144.1	147.8	150.6	161.6
164.7	167.4	172.6	185.9	190.8	195.4	202.0	209.8
209.7	212.7	215.3	228.7	233.9	236.5	241.2	260.1
262.7	268.3	274.9	296.4	298.9	305.3	310.9	332.0
331.5	329.2	327.3	351.1	348.7	346.9	354.6	387.1
386.5	386.6	396.1	423.7	420.8	417.8	429.5	464.6
451.6	454.5	468.8	503.7	499.3	504.4	506.7	534.8
525.5	518.2	519.8	560.5	553.0	552.4	560.4	594.0
594.7	594.0	601.4	634.1	629.4	621.0	632.0	663.5
657.4	657.1	668.1	699.7	700.7	700.8	703.4	759.9
756.5	755.7	766.6	804.1	809.0	802.7	807.4	847.1
839.8	835.4	844.7	874.7	869.5	863.6	876.9	916.4
914.3	1032.9	1038.2	1047.0	1057.9	1054.3	1061.0	1125.6
1145.4	1156.1	1188.7	1194.7	1229.5	1242.6	1263.1	1350.6
1388.7	1375.5	1366.7	1366.7	1367.9	1359.0	1371.7	1411.6
1446.6	1435.4	1436.6	0.0				

* CONSUMER PRICE INDEX (CPI): PERIOD AVERAGES (1991=100)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 147 last updated: 1996Q3
* one line of data represents two years

35.4	35.7	35.4	35.7	36.0	36.4	36.7	36.7
37.1	37.4	37.4	37.4	38.4	38.4	38.4	38.8
39.1	39.5	39.5	39.8	40.2	40.6	41.3	41.3
42.1	42.5	42.5	42.5	42.8	43.0	42.9	42.8
43.3	43.4	43.4	43.8	44.6	44.8	44.8	45.2
45.6	45.9	46.1	46.5	47.5	48.3	48.6	49.1
50.1	50.7	51.4	52.1	53.3	54.4	54.9	56.0
57.3	58.2	58.8	59.5	60.7	61.8	62.3	62.9
63.9	64.6	64.8	65.2	66.2	66.9	67.1	67.4
68.3	68.9	68.9	69.1	70.3	71.2	72.2	72.8
74.3	75.4	76.0	76.6	78.5	79.8	81.0	82.1
83.2	84.1	85.2	85.9	86.5	86.9	87.9	88.1
88.9	89.4	89.4	90.1	91.1	91.6	91.4	91.7
91.8	91.5	91.1	90.9	91.3	91.6	91.6	91.7
92.2	92.6	92.7	93.1	94.4	95.3	95.3	95.9
96.9	97.5	97.9	98.8	99.6	100.6	102.0	102.7
104.0	105.1	105.6	106.5	108.6	109.6	110.3	110.4
111.9	112.6	113.3	113.2	114.1	114.7	115.2	115.2
115.9	116.5	116.9	0.0				

* TREASURY BILL RATE (on 3-month loans, percent per annum)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 85 last updated: 1996Q3
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	5.40	5.40
5.22	5.02	5.37	5.08	4.82	4.44	4.06	4.00
3.64	3.57	3.71	4.00	4.22	5.01	5.82	7.42
8.05	8.30	7.41	7.41	9.40	10.70	11.92	10.24
9.56	8.53	8.40	6.77	5.44	5.23	5.88	5.95
5.79	5.73	5.76	5.38	5.65	5.62	4.53	4.35
4.07	3.66	3.84	3.87	3.57	2.99	3.18	3.40
3.15	3.15	3.96	4.20	5.30	6.23	6.49	7.12
7.63	8.17	8.25	8.25	8.25	8.25	8.25	8.27
8.50	8.50	8.70	7.57	6.98	6.55	5.95	5.40
5.10	4.95	4.98	5.17	5.08	4.43	4.27	3.83
3.40	3.42	3.29	0.0				

* GROSS DOM. PROD. 1990 PRICES (billions of Deutsche Mark)
* periodicity: QUARTERLY seasonally: ADJUSTED (at Annual Rates)
* starts: 1960Q1 stops: 1996Q4 no. of obs: 22 last updated: 1996Q2
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	2512.0	2561.7	2555.5	2560.9
2604.8	2587.1	2590.2	2580.5	2753.7	2761.8	2788.7	2790.3
2822.3	2846.6	2863.5	2892.1	2899.0	2919.4	2922.5	2922.5
2908.6	2951.8	0.0	0.0				

* GROSS NATIONAL PRODUCT: 1990 PRICES (billions of Deutsche Mark)
* periodicity: QUARTERLY seasonally: ADJUSTED (at Annual Rates)
* starts: 1960Q1 stops: 1994Q4 no. of obs: 89 last updated: 1994Q1
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1606.5	1610.0	1621.7	1656.1	1703.1	1701.9	1707.8	1714.6
1722.1	1718.6	1714.2	1675.0	1653.4	1652.6	1668.6	1701.6
1737.6	1757.2	1767.9	1780.3	1798.1	1801.6	1811.5	1831.9
1838.9	1868.3	1886.2	1908.4	1902.6	1945.8	1940.8	1971.5
1999.4	1963.1	1968.8	1963.1	1972.1	1956.6	1979.0	1972.1
1951.8	1962.4	1952.8	1949.2	1963.5	1997.0	1984.7	2014.0
2062.2	2020.2	2063.7	2080.3	2074.6	2085.8	2123.1	2117.7
2104.1	2156.7	2167.4	2168.4	2145.1	2168.0	2201.0	2216.9
2228.5	2235.4	2266.0	2287.5	2319.4	2321.2	2337.2	2358.6
2402.9	2427.9	2475.8	2499.0	2540.0	2543.2	2545.0	2533.6
2565.1	2567.8	2545.7	2527.0	2488.8	2500.3	2509.0	2513.6
2543.0	0.0	0.0	0.0				

(II) OECD MEI DATABASE

* GNP 1985 PRICES (billion DM)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 92Q4 no. of obs: 132 last updated: 92Q4
* one line of data represents two years

829.5	842.6	869.7	879.5	890.3	886.1	894.0	901.0
920.7	930.0	944.5	945.0	911.8	951.5	988.5	990.8
998.3	1024.5	1028.7	1043.2	1061.0	1079.2	1080.6	1094.2
1115.7	1118.0	1110.1	1098.8	1094.6	1104.4	1104.9	1134.8
1124.1	1156.8	1197.0	1217.6	1206.4	1248.5	1289.1	1304.1

1269.5	1333.6	1341.5	1358.3	1350.4	1361.1	1374.2	1375.6
1402.8	1411.2	1427.5	1449.5	1487.4	1481.8	1495.8	1492.5
1505.6	1491.6	1501.0	1470.1	1447.2	1461.7	1479.4	1493.9
1529.9	1555.7	1544.9	1580.0	1587.5	1583.7	1579.5	1624.4
1623.9	1640.8	1644.0	1674.4	1679.1	1710.9	1718.4	1735.2
1759.1	1728.7	1734.3	1723.1	1739.0	1726.8	1743.2	1736.2
1724.9	1726.3	1714.2	1712.8	1737.1	1753.0	1744.1	1775.0
1817.5	1777.3	1812.8	1830.6	1819.4	1832.0	1865.7	1861.5
1846.5	1892.3	1905.9	1905.4	1890.7	1904.5	1933.7	1946.3
1967.9	1971.1	1977.2	1999.6	2041.2	2030.4	2050.0	2070.6
2114.4	2132.5	2174.9	2195.3	2231.3	2234.0	2235.2	2225.7
2256.8	2256.8	2234.0	2220.1				

* GNP IMPLICIT PRICE LEVEL: 1985=100
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 92Q4 no. of obs: 132 last updated: 92Q4
* one line of data represents two years

35.1	34.8	36.0	35.8	36.4	36.7	37.4	37.9
37.9	38.5	38.8	39.1	39.5	39.8	39.8	40.1
40.5	40.7	41.2	41.6	42.0	42.2	42.7	43.0
43.5	43.6	44.3	44.1	44.4	44.4	44.5	44.6
45.2	45.2	45.3	46.1	46.7	46.9	47.3	48.6
49.8	50.9	51.1	52.1	53.8	54.9	55.3	56.2
57.0	57.7	58.1	59.2	60.5	61.1	61.6	63.6
63.7	65.4	66.6	68.6	68.8	69.8	70.1	71.3
71.4	72.3	73.3	73.1	74.1	75.4	75.3	76.2
77.4	78.3	79.0	79.3	80.4	80.9	82.2	83.0
84.0	85.5	86.0	86.8	87.3	88.5	89.3	90.9
91.7	92.3	93.4	94.1	95.0	95.3	96.4	97.0
97.2	97.5	98.0	98.6	99.1	99.5	100.4	101.1
102.0	102.7	103.6	104.2	105.0	105.1	105.0	105.7
105.9	106.7	107.1	107.7	108.7	109.2	109.7	111.0
112.2	112.9	114.1	114.5	115.7	118.1	118.9	119.8
121.5	123.2	124.3	125.2				

* MONEY SUPPLY M1 (billion DM, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1993Q4 no. of obs: 135 last updated: 1993Q3
* one line of data represents two years

46.409	47.770	48.387	51.072	49.465	51.960	53.902	58.707
55.273	57.858	59.251	63.351	59.116	62.160	63.553	67.760
64.439	67.860	68.536	73.037	70.276	74.196	74.353	78.520
74.316	77.908	76.933	79.614	75.821	78.950	80.406	87.921
81.128	85.487	86.963	93.466	87.551	92.759	94.281	99.429
93.452	98.550	99.142	108.219	101.879	110.146	112.389	121.522
117.175	125.996	128.264	139.298	133.354	133.359	129.098	142.862
132.873	140.795	141.477	158.432	149.341	160.641	164.425	179.898
166.800	180.171	176.915	186.852	179.738	190.483	193.081	208.076
204.242	215.359	217.498	237.909	225.460	233.116	230.226	247.869
228.689	237.063	237.761	257.335	232.375	242.478	234.283	255.277
237.602	250.712	248.466	273.047	263.403	277.861	273.975	295.795
272.542	282.798	281.455	314.235	285.223	294.411	297.800	334.097
313.436	329.320	326.936	358.747	336.807	358.731	357.178	385.170
369.503	393.475	389.095	426.997	403.190	412.016	408.688	450.623
412.827	483.000	502.800	584.300	530.300	541.000	546.800	604.000

556.600 576.400 588.200 669.600 610.100 631.500 643.700 0.000

* M1 + QUASI-MONEY (billion DM, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 1960Q1
* stops: 1993Q4 no. of obs: 135 last updated: 1993Q3
* one line of data represents two years

67.078	68.570	69.617	72.692	72.186	75.437	77.194	82.172
79.900	82.976	83.851	88.544	84.543	87.730	89.275	94.508
91.025	94.694	95.113	100.813	97.819	101.413	100.991	106.916
103.279	107.615	108.483	113.384	110.561	114.191	116.322	127.574
123.030	129.414	133.033	142.327	137.282	145.455	150.538	156.875
148.834	158.950	160.293	173.383	169.345	179.487	181.813	198.595
192.068	204.661	210.432	232.330	231.316	241.730	246.920	265.861
255.762	262.227	258.004	279.603	260.153	254.890	254.176	279.318
260.571	274.580	276.006	298.180	283.089	292.816	298.432	331.812
310.983	324.448	335.710	375.408	350.342	367.051	369.228	406.492
389.602	399.451	401.309	440.616	421.508	443.780	451.705	478.132
460.014	473.352	477.318	502.200	478.033	485.028	487.359	515.364
485.729	501.743	509.787	542.322	511.413	528.211	527.053	577.103
547.989	562.937	570.222	610.925	584.824	603.517	602.161	645.646
616.813	639.419	645.226	696.075	674.102	691.758	705.639	776.381
752.944	839.900	881.000	987.600	956.100	976.700	999.40	1084.500
1060.10	1091.30	1139.30	1196.500	1162.20	1193.10	1206.40	0.000

* MONEY SUPPLY M1 (billion DM, end of period)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 148 last updated: 1996Q4
* one line of data represents two years

47.2	47.8	48.3	49.0	50.4	52.0	53.8	56.3
56.4	57.8	59.2	60.8	60.4	62.0	63.5	65.1
65.9	67.6	68.6	70.2	72.1	73.8	74.5	75.4
76.2	77.4	76.9	76.4	77.9	78.5	80.6	84.0
83.5	84.9	87.0	89.2	89.5	91.5	94.6	94.4
95.9	97.5	99.8	102.8	104.8	109.3	113.5	115.4
120.8	125.3	129.2	131.8	136.8	132.2	130.1	134.4
136.1	139.6	141.9	149.8	153.0	158.5	165.8	170.6
171.1	179.1	179.6	177.7	184.8	189.9	196.2	197.2
210.1	214.8	220.3	225.5	231.1	231.8	233.3	233.6
232.9	234.5	240.7	244.1	237.7	240.9	238.0	242.2
243.8	249.7	252.9	258.7	270.9	277.1	278.8	278.9
279.3	280.9	285.3	294.3	292.1	292.3	300.2	314.1
320.2	324.4	331.2	338.8	344.3	355.6	363.7	364.1
378.9	392.1	396.9	402.9	414.0	411.1	415.8	423.9
423.1	423.5	436.7	453.1	541.9	539.4	552.1	559.7
567.7	576.6	597.8	626.5	624.6	634.2	653.9	676.6
691.9	707.2	713.6	711.8	712.9	724.9	738.7	757.6
788.1	800.5	814.9	843.1				

* IMPORTS C.I.F. (billion DM, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 148 last updated: 1996Q4
* one line of data represents two years

3.47	3.56	3.50	3.69	3.55	3.68	3.70	3.88
4.11	4.05	4.10	4.28	4.17	4.45	4.52	4.35

4.58	4.66	5.00	5.32	5.58	5.72	5.99	6.06
6.13	6.02	5.99	6.07	5.62	5.64	5.81	6.38
6.41	6.46	7.00	7.23	7.78	8.16	8.37	8.40
8.87	8.95	9.25	9.48	9.87	9.94	10.15	9.92
10.17	10.54	10.79	11.41	11.89	11.97	11.81	12.86
13.84	15.01	15.88	15.50	14.51	15.09	15.51	16.52
17.22	18.03	19.12	19.07	19.07	19.29	19.88	19.93
19.59	19.89	20.53	21.26	21.84	23.67	25.48	26.60
28.29	28.75	28.07	28.69	30.27	30.39	31.53	30.91
31.86	31.55	30.92	30.82	30.53	31.83	32.92	34.46
35.12	35.35	37.10	37.40	39.50	38.76	39.15	37.97
36.48	34.87	33.37	33.53	33.10	34.23	34.50	34.63
33.63	35.60	37.43	39.17	40.27	42.50	42.43	43.70
43.57	43.57	47.87	51.70	53.47	54.03	54.97	53.37
54.47	52.77	52.37	51.47	46.70	46.57	47.73	46.83
48.60	51.33	52.77	53.60	54.13	55.53	56.20	55.30
56.17	56.50	56.87	59.30				

* EXPORTS F.O.B. (billion DM, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 1960Q1
* stops: 1996Q4 no. of obs: 148 last updated: 1996Q4
* one line of data represents two years

3.95	3.84	3.94	4.24	4.21	4.24	4.34	4.26
4.26	4.38	4.53	4.52	4.38	4.86	5.06	5.20
5.36	5.34	5.33	5.55	5.57	5.76	5.94	6.14
6.37	6.55	6.85	6.72	7.06	7.19	7.21	7.63
7.83	7.73	8.49	9.17	8.69	9.62	9.77	9.80
9.94	10.28	10.65	10.84	11.13	11.25	11.71	11.06
11.71	12.05	12.37	13.52	13.86	14.52	15.31	15.95
18.40	19.24	19.96	19.70	18.34	18.30	18.36	19.20
20.16	20.68	22.04	21.81	22.10	22.58	22.72	23.45
22.81	23.27	24.03	24.88	24.54	26.09	26.96	27.52
29.40	29.38	28.81	29.17	30.57	32.63	34.19	34.78
35.98	35.94	35.20	34.89	34.86	35.24	36.04	37.43
39.25	38.18	41.79	43.70	44.67	45.12	45.53	44.76
44.98	44.23	43.76	43.23	42.57	43.80	44.07	45.20
42.17	47.03	48.17	50.97	52.23	53.83	54.20	53.60
55.53	53.23	56.63	56.40	55.93	54.07	56.60	56.93
56.73	55.50	56.57	53.67	51.70	51.33	51.83	53.07
54.73	58.20	58.37	59.20	61.10	62.60	62.97	62.83
63.33	63.73	65.97	68.43				

(I) IMF IFS DATABASE

* MONEY (Billions of Francs: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 146 last updated: 96Q2
* one line of data represents two years

97.91	102.24	106.83	113.34	115.34	119.61	123.03	130.94
132.41	140.25	146.14	154.71	156.88	163.29	169.43	177.16
175.72	179.91	185.85	191.83	189.31	196.62	201.55	209.86
207.84	215.44	218.86	226.13	223.00	230.66	230.85	236.93
226.12	245.71	248.20	255.86	252.71	259.02	259.73	250.53
241.73	249.67	254.15	277.71	274.46	289.09	292.86	310.48
307.65	326.67	337.10	357.23	339.48	361.79	363.04	391.74
402.41	404.05	392.59	451.27	417.95	441.37	455.98	507.97
488.79	514.51	513.09	545.95	524.65	542.98	555.58	606.67
579.90	599.10	620.30	665.90	660.30	691.90	697.80	751.40
713.80	742.60	757.90	801.20	784.70	822.00	858.20	899.80
881.60	925.10	942.90	985.00	965.30	1014.90	1015.50	1108.40
1073.90	1119.50	1136.40	1219.20	1193.10	1215.60	1226.60	1312.40
1310.50	1320.30	1337.30	1406.80	1359.90	1401.20	1377.80	1471.10
1366.50	1416.00	1400.40	1531.60	1418.00	1476.00	1511.00	1633.00
1523.00	1581.00	1566.00	1703.00	1522.00	1576.00	1532.00	1609.00
1487.00	1540.00	1520.00	1608.00	1498.00	1544.00	1505.00	1620.00
1540.00	1570.00	1542.00	1673.00	1530.00	1588.00	1591.00	1819.00
1649.00	1701.00	0.00	0.00				

* QUASI-MONEY (Billions of Francs: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 146 last updated: 96Q2
* one line of data represents two years

16.40	17.54	19.20	20.81	23.57	25.30	26.46	27.47
30.19	30.69	32.27	33.66	36.22	35.77	36.49	37.27
40.47	41.57	43.47	45.20	48.55	49.96	52.91	54.75
60.48	62.71	66.61	69.96	75.80	79.37	98.07	109.01
123.91	124.31	130.84	135.21	145.94	149.52	159.96	166.85
184.13	189.78	200.09	212.19	231.63	241.54	258.23	280.72
301.57	318.71	335.76	352.48	367.93	382.41	410.72	431.83
459.42	469.77	503.37	525.30	563.96	573.72	603.17	630.84
673.24	694.25	720.05	747.65	793.08	804.04	844.97	886.82
924.60	942.20	967.10	1015.20	1058.00	1073.70	1110.70	1171.10
1208.20	1212.60	1240.20	1300.40	1346.00	1368.40	1365.50	1422.90
1486.90	1506.50	1532.50	1592.70	1644.00	1646.80	1659.80	1752.70
1790.70	1804.70	1827.60	1888.00	1931.00	1924.10	1951.80	2005.90
2103.90	2088.50	2110.40	2164.50	2219.10	2238.40	2271.50	2333.40
2414.40	2413.88	2453.80	2484.10	2417.00	2435.00	2464.00	2479.00
2527.00	2476.00	2484.00	2509.00	2556.00	2524.00	2542.00	2564.00
2594.00	2581.00	2573.00	2618.00	2665.00	2652.00	2659.00	2766.00
2814.00	2816.00	2872.00	3013.00	3144.00	3196.00	3245.00	3374.00
3420.00	3420.00	0.00	0.00				

* M2 (Billions of Francs: Average of Figures for Last Month in
*   Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 63 last updated: 93Q2
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	1070.6
1067.2	1093.0	1132.4	1213.3	1232.0	1271.1	1290.0	1371.4
1343.0	1370.5	1399.4	1482.4	1482.1	1523.4	1581.1	1685.1
1691.8	1743.8	1787.0	1887.2	1891.3	1951.3	1963.1	2138.5
2114.1	2163.2	2196.2	2335.9	2301.4	2321.6	2332.9	2473.4
2479.6	2469.1	2480.0	2587.7	2523.1	2558.1	2551.8	2699.1
2593.8	2636.3	2641.7	2798.9	2699.1	2742.6	2784.0	2930.6
2768.3	2819.0	2811.0	2956.0	2770.0	2800.0	2756.0	2858.0
2710.0	2737.0	2711.0	2821.0	2686.0	2735.0	0.0	0.0

* M3 (Billions of Francs: Average of Figures for Last Month in
*   Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 63 last updated: 93Q2
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	1405.2
1418.1	1450.9	1491.8	1579.4	1607.9	1651.5	1691.1	1800.9
1796.4	1830.6	1872.3	1973.5	2002.0	2060.2	2092.1	2192.8
2241.1	2311.2	2359.6	2463.8	2497.6	2560.8	2593.3	2787.3
2805.9	2888.0	2949.7	3092.6	3112.2	3150.9	3185.0	3320.9
3378.1	3398.8	3446.5	3543.6	3567.4	3637.5	3703.9	3881.9
3874.6	3948.7	4035.7	4208.6	4237.0	4334.0	4444.0	4624.0
4623.0	4737.0	4806.0	5034.0	4993.0	5043.0	5067.0	5160.0
5190.0	5295.0	5308.0	5430.0	5465.0	5445.0	0.0	0.0

* CONSUMER PRICE INDEX (CPI): PERIOD AVERAGES (1990=100)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 147 last updated: 96Q3
* one line of data represents two years

14.5	14.5	14.6	14.8	14.8	14.8	14.9	15.3
15.5	15.6	15.7	16.0	16.3	16.4	16.7	16.8
16.9	16.9	17.0	17.2	17.3	17.5	17.5	17.6

17.8	17.8	18.0	18.1	18.3	18.3	18.5	18.7
19.0	19.1	19.3	19.7	20.0	20.3	20.5	20.8
21.3	21.5	21.7	22.0	22.3	22.6	22.9	23.3
23.6	23.9	24.4	24.9	25.1	25.6	26.2	26.9
27.9	29.1	30.0	31.0	31.8	32.6	33.3	34.0
34.9	35.7	36.5	37.4	38.0	39.2	40.1	40.9
41.5	42.7	43.8	44.7	45.7	47.0	48.6	49.9
51.8	53.5	55.1	56.7	58.4	60.3	62.7	64.7
66.6	68.6	69.5	70.9	72.7	74.8	76.4	77.8
79.1	80.6	82.0	83.1	84.2	85.8	86.6	87.1
87.2	87.8	88.4	88.9	90.1	90.8	91.4	91.8
92.2	93.2	94.0	94.5	95.4	96.5	97.2	98.0
98.7	99.6	100.4	101.4	102.1	102.9	103.5	104.4
104.9	105.7	105.7	106.3	107.1	107.8	108.0	108.5
109.0	109.6	109.8	110.3	110.9	111.4	111.8	112.4
113.2	114.1	113.8	0.0				

* TREASURY BILL RATE (Percent Per Annum)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 71 last updated: 96Q3
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	6.85	7.76	10.82	12.50
13.00	12.65	11.68	11.47	11.98	15.94	17.44	15.69
17.39	17.33	14.71	13.35	13.02	13.10	12.93	12.53
12.68	12.98	12.05	10.44	11.08	10.52	10.05	9.05
8.78	7.02	7.20	7.70	8.33	8.13	7.92	8.51
7.94	7.75	7.66	8.16	8.97	8.83	9.15	10.40
10.63	9.91	10.14	10.05	9.83	9.58	9.69	9.65
10.24	10.32	10.65	10.73	11.62	8.09	7.34	6.58
6.26	5.76	5.63	5.49	6.51	7.59	6.17	6.06
4.39	3.87	3.76	0.00				

* GROSS DOM. PROD. 1990 PRICES (Billions of Francs)
* periodicity: QUARTERLY seasonally: ADJUSTED (at Annual Rates)
* starts: 60Q1 stops: 96Q4 no. of obs: 86 last updated: 96Q2
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	3720.7	3769.8	3758.7	3875.9
3945.6	3985.7	4029.4	4062.6	4121.2	4108.6	4142.5	4147.8
4800.2	4880.6	4889.5	4958.9	4968.1	5013.7	5095.9	5114.1

5129.5	5087.5	5093.1	5080.2	5050.6	5101.7	5132.8	5180.3
5193.2	5227.8	5218.6	5242.4	5242.4	5264.3	5246.1	5280.7
5478.1	5528.9	5492.1	5507.6	5506.6	5535.3	5583.1	5608.1
5622.3	5695.0	5709.0	5729.6	5734.7	5803.8	5854.4	5913.9
6002.8	6039.6	6094.5	6150.5	6278.9	6331.7	6360.4	6393.9
6480.8	6502.6	6546.6	6510.5	6510.4	6542.0	6580.9	6611.0
6658.8	6655.1	6657.3	6624.3	6535.4	6536.1	6553.8	6567.7
6606.6	6717.5	6770.4	6839.5	6877.7	6887.2	6898.2	6868.1
6945.2	6921.0	0.0	0.0				

(II) OECD MEI DATABASE

* GDP CONSTANT PRICES OF 1980 (Billions of FFrancs)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 108 last updated: 96Q4
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1980.2	2014.8	2047.1	2073.6	2079.2	2112.3	2147.2	2165.1
2193.6	2202.8	2225.1	2259.1	2300.7	2327.7	2353.1	2382.3
2398.8	2414.7	2440.0	2401.4	2385.4	2395.8	2406.5	2440.3
2475.3	2503.4	2517.8	2540.1	2572.3	2579.4	2597.8	2609.9
2637.4	2678.9	2680.9	2709.4	2727.6	2751.6	2784.2	2790.3
2812.7	2801.9	2809.0	2809.6	2793.2	2831.5	2856.5	2884.0
2891.2	2915.4	2913.2	2934.9	2925.3	2938.8	2927.6	2943.8
2976.7	2956.3	2970.0	2986.8	2988.4	3017.2	3044.0	3063.9
3062.6	3107.1	3120.3	3130.4	3123.1	3154.6	3191.1	3239.4
3272.5	3282.3	3324.2	3358.5	3406.8	3430.2	3466.9	3488.7
3531.7	3528.9	3570.6	3556.4	3550.4	3567.1	3590.1	3601.0
3628.8	3621.8	3620.9	3615.7	3563.4	3564.1	3571.4	3576.2
3598.5	3658.2	3687.9	3725.7	3745.1	3750.7	3758.4	3739.8
3786.4	3781.5	3810.7	3816.8				

* IMPLICIT PRICE LEVEL: 1990=100
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 107 last updated: 96Q3
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20.8	21.1	21.5	21.8	22.2	22.5	22.8	23.2
23.5	24.0	24.5	25.0	25.3	25.9	26.6	27.4
27.8	28.9	30.0	31.0	32.1	32.8	33.7	34.4
35.3	36.5	37.5	38.5	38.9	39.8	40.9	41.8
42.6	43.7	45.2	46.2	47.2	48.2	49.4	50.9
52.2	53.7	55.5	56.7	58.0	59.4	61.4	63.9
65.7	67.6	68.4	69.5	71.9	73.8	75.3	76.8
78.2	79.5	80.6	81.6	82.7	83.9	85.3	86.4

88.0	88.7	89.4	89.7	90.5	91.1	91.5	92.5
93.1	93.6	94.7	95.6	96.2	96.8	97.4	98.9
99.0	100.2	100.4	101.0	101.9	102.6	103.2	104.0
104.8	105.2	105.7	106.2	107.4	108.0	108.4	108.6
109.2	109.5	109.9	110.3	110.9	111.4	111.8	112.2
112.8	113.2	113.4	0.0				

* MONEY SUPPLY M1 (billion francs, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 95 last updated: 93Q2
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	208.04
201.61	208.38	212.33	232.20	229.22	241.27	244.47	259.56
257.28	273.14	281.79	298.86	284.03	302.14	303.47	327.70
316.85	328.51	328.43	365.76	349.50	369.55	382.47	426.47
409.37	430.41	429.42	457.42	438.89	454.51	465.57	603.20
582.04	601.25	622.08	668.81	662.89	694.59	700.75	754.82
717.51	746.41	761.48	806.04	789.00	826.24	864.18	905.80
886.63	930.96	949.25	992.66	972.92	1024.77	1025.40	1121.08
1087.52	1136.28	1156.43	1245.73	1211.75	1241.17	1238.40	1328.69
1331.04	1337.22	1339.20	1408.36	1362.81	1404.35	1376.45	1471.09
1366.49	1392.50	1383.90	1508.60	1415.70	1473.90	1507.60	1634.60
1514.80	1579.70	1565.30	1688.50	1514.80	1566.10	1522.60	1609.20
1487.60	1539.10	1518.80	1607.10	1504.60	1547.50	0.00	0.00

* M1 + QUASI-MONEY (billion francs, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 95 last updated: 93Q2
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	300.00
300.60	310.27	319.47	345.91	352.81	369.71	381.34	408.08
416.69	442.44	459.20	484.65	478.50	503.69	519.47	555.66
558.88	576.13	593.90	642.59	646.57	671.77	700.23	758.81
764.60	795.57	809.01	852.23	858.48	878.85	913.25	1406.94
1419.94	1452.81	1493.54	1582.00	1610.41	1654.20	1693.96	1804.56
1800.32	1834.03	1875.37	1978.27	2006.75	2064.88	2097.68	2194.22
2236.50	2300.83	2345.37	2448.09	2473.28	2530.08	2544.16	2729.21
2725.82	2791.41	2839.47	2990.16	3000.22	3023.61	3059.57	3167.51
3190.22	3187.95	3201.68	3317.90	3353.90	3436.05	3456.52	3616.32
3602.91	3948.70	4035.70	4208.60	4241.60	4339.20	4447.90	4627.60
4626.30	4736.60	4806.00	5034.30	4992.50	5042.90	5066.80	5160.30
5189.60	5295.00	5307.60	5429.50	5474.20	5465.70	0.00	0.00

* MONEY SUPPLY M1 (billion francs, end of period)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 148 last updated: 96Q4
* one line of data represents two years

82.0	85.3	89.3	92.7	96.7	99.8	102.7	107.1
111.1	117.0	121.8	126.5	131.7	136.2	141.0	144.7
147.7	150.4	153.8	156.7	160.1	164.1	168.4	169.2
174.3	175.8	178.7	182.1	186.9	188.3	188.6	190.5
189.3	200.7	202.9	205.4	209.7	209.1	210.3	200.3
203.5	205.6	212.2	223.0	231.2	237.9	244.4	248.3
259.2	269.2	281.5	284.8	286.0	297.6	303.4	310.8
318.8	323.5	328.4	345.7	351.9	364.2	382.6	402.1
412.3	424.4	429.3	431.6	442.0	448.6	464.8	579.5
584.5	594.9	619.0	642.6	666.0	687.4	697.5	725.5
721.8	738.7	758.0	775.3	795.3	817.8	861.3	869.8
894.0	921.0	946.9	951.1	980.4	1013.6	1024.0	1071.2
1094.1	1123.0	1156.4	1187.1	1216.9	1226.2	1240.3	1264.4
1333.9	1318.8	1342.9	1338.4	1365.4	1389.3	1383.1	1395.2
1383.8	1386.0	1388.0	1423.6	1443.0	1468.9	1509.0	1540.0
1544.7	1567.0	1566.5	1588.4	1544.4	1550.6	1525.5	1516.3
1514.8	1516.0	1524.0	1516.6	1521.0	1523.2	1515.5	1532.6
1563.4	1551.8	1553.4	1568.6	1545.4	1564.9	1602.4	1686.4
1665.0	1676.3	1682.6	1696.2				

* IMPORTS F.O.B. (billion francs, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 148 last updated: 96Q4
* one line of data represents two years

2.51	2.48	2.68	2.65	2.59	2.77	2.76	2.90
3.04	2.96	3.21	3.33	3.24	3.59	3.65	3.88
4.06	4.22	4.12	4.11	4.05	4.19	4.30	4.47
4.67	4.87	4.98	5.06	5.21	4.98	5.04	5.20
5.42	4.90	6.16	6.47	6.74	7.32	7.56	7.95
7.92	8.13	8.47	8.76	8.61	8.78	9.46	9.65
10.01	10.39	10.51	11.55	11.85	12.40	13.37	14.21
18.20	20.34	21.27	19.77	18.64	17.33	17.97	19.87
21.39	23.38	25.66	27.46	27.39	27.27	27.69	27.99
29.01	29.11	29.40	30.83	32.21	34.87	38.73	40.20
44.46	45.58	46.01	47.62	47.98	50.79	54.17	56.14
56.51	59.76	62.98	62.88	62.64	63.26	64.53	66.75
70.84	72.50	72.45	75.81	78.82	78.41	77.96	77.06
72.68	72.78	72.47	70.17	73.05	75.60	77.53	80.67
80.42	83.45	88.84	90.74	95.99	99.18	98.69	101.91
101.43	100.93	102.92	103.66	103.09	103.41	106.97	103.55
103.37	102.91	100.92	98.77	91.47	91.66	91.92	92.57
97.92	100.82	102.92	106.82	111.08	112.93	113.55	112.10
113.92	113.36	115.16	116.97				

* EXPORTS F.O.B. (billion francs, monthly averages)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 148 last updated: 96Q4
* one line of data represents two years

2.91	2.74	2.77	2.86	2.87	2.95	3.07	3.04
3.05	2.94	3.00	2.98	3.09	3.40	3.38	3.46
3.63	3.65	3.63	3.84	3.89	4.05	4.26	4.38
4.35	4.54	4.59	4.51	4.51	4.68	4.72	4.89
5.19	4.26	5.83	5.61	5.86	6.20	6.48	7.12
8.10	8.16	8.45	8.71	9.16	9.16	9.94	10.16

10.16	11.08	11.16	12.33	12.52	13.10	13.88	14.65
17.21	18.47	19.33	19.36	19.25	18.43	18.73	19.29
20.69	22.41	23.34	24.64	25.71	26.01	26.93	27.94
28.94	29.45	29.79	30.96	32.60	34.64	36.98	38.36
39.68	40.69	40.59	42.57	44.22	47.84	50.11	50.02
51.29	51.34	52.87	55.43	55.18	58.97	61.52	65.15
67.61	69.83	71.98	73.91	75.26	76.65	74.88	75.30
73.61	70.94	72.39	70.98	70.70	71.54	75.47	78.62
78.63	82.25	84.89	86.77	94.32	94.57	94.50	97.65
100.20	97.02	96.99	98.19	98.33	100.45	104.00	104.38
105.16	106.51	102.93	101.69	95.92	98.42	100.24	102.42
103.62	107.24	109.49	115.37	116.83	119.65	116.60	118.98
120.47	120.07	124.05	125.84				

(I) IMF IFS DATABASE

* MONEY (M1) (Trillions of Lire: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 146 last updated: 96Q2
* one line of data represents two years

6.287	6.531	6.915	7.414	7.244	7.492	7.795	8.601
10.529	10.863	11.405	12.180	12.210	12.449	12.775	9.423
9.095	9.438	9.772	10.945	10.595	11.079	11.489	12.672
12.269	12.653	13.060	14.154	13.921	14.476	15.004	16.428
15.992	16.794	17.386	18.970	18.953	19.709	20.449	22.477
22.833	24.929	26.302	29.885	29.417	31.003	32.131	35.567
35.157	36.696	37.699	42.229	41.680	45.151	46.825	52.460
51.765	53.889	53.946	57.194	54.971	54.897	57.600	64.011
65.681	67.192	68.682	76.311	77.161	79.953	82.625	92.702
95.067	99.764	103.932	116.763	116.832	122.098	127.440	145.069
140.635	143.250	143.598	164.616	159.843	159.653	158.374	181.044
171.378	175.575	180.620	212.029	198.824	202.631	211.133	239.210
226.247	228.723	234.391	269.575	258.007	259.709	267.923	297.951
284.623	287.444	294.277	331.039	312.915	321.552	325.879	357.151
335.494	343.327	352.666	386.037	361.559	374.275	377.967	433.334
409.340	420.230	423.720	482.230	443.700	455.810	459.910	537.530
486.690	499.310	493.670	545.790	505.090	517.630	522.110	579.080
561.690	563.090	549.110	597.330	551.840	557.150	553.020	604.490
548.420	560.370	0.000	0.000				

* QUASI-MONEY (Trillions of Lire: End of Period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 94Q4 no. of obs: 139 last updated: 94Q3
* one line of data represents two years

3.794	3.887	4.011	4.252	4.376	4.485	4.648	4.928
5.092	5.211	5.416	5.795	5.959	6.060	6.218	8.777
8.871	8.898	9.058	9.650	9.853	9.992	10.279	11.075
11.322	11.504	11.852	12.646	12.844	13.041	13.337	14.113
14.318	14.512	14.821	15.656	15.764	15.815	16.091	16.750
16.543	16.183	15.955	16.481	16.444	16.865	17.552	18.883
19.518	20.189	20.972	22.552	23.307	23.903	24.761	27.393
28.947	29.839	30.675	33.589	37.418	40.378	42.529	46.661
48.653	49.508	51.371	56.951	58.993	60.251	61.996	69.485
71.608	73.134	75.363	83.085	84.573	85.983	88.294	97.082
96.614	96.425	97.999	108.617	107.663	106.493	107.210	119.195
118.845	119.255	122.646	142.748	140.738	142.054	145.640	159.286
158.435	158.911	163.679	178.021	181.953	183.628	186.452	199.032
196.243	194.672	198.616	214.519	219.581	220.573	221.267	234.644
234.532	237.785	244.409	257.162	267.603	270.067	274.497	285.361
292.110	295.140	303.570	323.730	323.740	324.930	329.180	345.800
350.860	357.090	366.560	388.240	399.780	402.420	409.400	423.980
431.030	425.180	419.360	0.000				

* CONSUMER PRICE INDEX (CPI): PERIOD AVERAGES (1990=100)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 147 last updated: 96Q3
* one line of data represents two years

7.3	7.3	7.4	7.4	7.4	7.5	7.5	7.5
7.7	7.8	7.8	8.0	8.3	8.3	8.4	8.6
8.7	8.8	9.0	9.1	9.2	9.2	9.3	9.4
9.5	9.5	9.5	9.6	9.8	9.8	9.9	10.0
10.0	10.0	10.0	10.0	10.1	10.2	10.3	10.4
10.6	10.7	10.8	11.0	11.1	11.2	11.3	11.5
11.6	11.8	12.0	12.3	12.7	13.1	13.4	13.8
14.5	15.2	16.2	17.1	17.8	18.2	18.6	19.1
19.9	21.2	21.5	22.9	24.0	24.9	25.5	26.4
27.1	27.9	28.6	29.4	30.6	31.7	32.8	34.6
36.9	38.3	39.9	42.0	43.6	45.5	46.9	49.1
51.0	52.6	54.8	57.2	59.3	61.0	62.4	64.6
67.4	68.9	69.9	71.6	73.7	75.4	76.3	78.0
79.2	80.0	80.5	81.5	82.6	83.6	84.4	85.8
86.9	87.8	88.6	90.2	92.0	93.5	94.3	95.9
97.9	99.1	100.5	102.3	104.2	105.7	106.9	108.4
110.0	111.3	112.2	113.5	114.8	116.2	117.3	118.5
120.0	120.8	121.8	123.1	125.2	127.3	128.6	130.0
131.3	132.9	133.2	0.0				

* TREASURY BILL RATE (WGHTD AV BEFORE TAX, Percent Per Annum)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 78 last updated: 96Q3
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	16.23	13.82	12.89
12.28	11.97	11.87	11.85	12.05	11.96	12.06	13.98
15.24	15.74	15.80	16.88	17.55	19.45	20.37	21.43
20.24	19.95	18.69	18.88	18.95	17.78	17.46	17.36
16.11	15.25	15.16	14.94	13.53	14.17	13.95	13.20
13.35	11.40	10.58	10.29	10.00	9.88	11.20	11.85
10.81	10.78	11.36	11.55	12.12	11.97	12.80	13.43
12.93	12.12	11.77	12.69	13.01	12.06	12.44	12.63
12.44	13.53	16.13	15.17	12.31	11.46	9.62	8.94
8.78	8.52	9.50	9.86	10.68	11.14	10.81	10.77
9.70	8.69	8.41	0.00				

* GROSS DOM. PROD. 1990 PRICES (Trillions of Lire)
* periodicity: QUARTERLY seasonally: ADJUSTED (at Annual Rates)
* starts: 1960Q1 stops: 1993Q4 no. of obs: 83
* last updated: 1993Q3
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	639.72	704.96	720.61	786.06
717.46	721.09	709.51	691.28	691.57	684.98	682.61	700.56
778.43	792.69	796.34	816.32	822.69	808.69	807.71	808.56
821.38	824.99	831.26	853.72	864.37	858.51	868.73	903.07
918.19	912.34	895.40	904.99	915.57	910.47	904.45	906.78
923.17	902.92	901.01	891.29	899.35	891.75	901.26	910.86
1077.86	1097.01	1113.57	1111.55	1124.71	1133.37	1139.62	1149.46
1148.97	1161.36	1162.50	1173.49	1178.60	1192.36	1198.77	1218.48
1240.95	1248.97	1261.68	1272.22	1268.21	1280.05	1291.09	1296.58
1307.23	1306.23	1320.22	1314.58	1320.03	1327.09	1330.16	1336.72
1343.04	1347.27	1339.90	1333.12	1330.40	1339.80	1333.40	0.00

(II) OECD MEI DATABASE

* GROSS DOMESTIC PRODUCT: 1990 PRICES (Trillions of Lire)
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 103 last updated: 96Q3
* one line of data represents two years

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	771.24	773.02	782.88	791.38
800.28	792.37	799.32	821.39	821.08	849.40	876.42	892.41
901.15	905.18	900.13	875.28	854.35	858.50	859.10	879.53
888.18	905.30	922.00	938.60	939.56	934.20	926.41	923.35
932.91	952.74	961.89	975.98	981.46	988.04	1009.42	1032.09
1045.19	1046.79	1037.15	1038.80	1045.78	1056.76	1049.48	1055.87
1059.10	1058.85	1049.34	1054.29	1059.82	1060.73	1069.25	1079.31
1090.41	1095.85	1101.34	1110.12	1114.16	1126.19	1133.85	1139.04
1144.79	1161.89	1173.03	1167.46	1170.93	1187.12	1193.13	1212.05
1225.66	1234.03	1244.74	1255.53	1266.32	1275.75	1285.85	1290.58
1301.05	1300.49	1313.68	1307.70	1314.32	1321.19	1323.06	1329.62
1335.84	1338.04	1325.22	1324.97	1317.60	1317.60	1305.11	1320.37
1327.00	1344.00	1354.00	1359.00	1380.00	1382.00	1390.00	1391.00
1397.00	1392.00	1400.00	0.00				

* IMPLICIT PRICE LEVEL: 1990=100
* periodicity: QUARTERLY seasonally: ADJUSTED starts: 60Q1
* stops: 96Q4 no. of obs: 99 last updated: 96Q3
* one line of data represents two years

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

10.4	10.8	10.8	11.0	11.3	11.7	12.3	12.6
13.1	13.8	14.6	15.3	16.1	16.4	16.8	17.2
17.9	19.1	20.1	21.2	22.4	22.9	23.7	24.5
25.4	26.0	27.1	27.9	28.8	30.1	31.3	33.0
34.6	36.4	37.8	39.4	40.9	42.7	45.0	47.0
48.6	50.3	52.4	54.4	56.5	58.1	59.7	62.0
64.0	65.4	66.4	67.6	69.6	70.9	72.6	73.7
75.4	76.6	77.9	78.7	79.9	81.7	82.1	83.3
84.6	86.1	87.8	88.9	90.3	91.9	93.4	95.2
97.4	98.8	100.3	103.1	104.8	106.5	108.4	109.7
110.9	111.9	112.8	114.0	115.5	116.8	118.2	118.9
120.4	121.1	122.0	123.2	124.8	127.0	128.8	130.6
132.0	132.9	133.9	0.0				

* MONEY SUPPLY M1 (100 billion Lire, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 134 last updated: 93Q2
* one line of data represents two years

63.86	66.26	70.22	75.16	73.55	75.99	79.26	87.30
85.31	88.65	92.69	101.74	101.82	105.09	108.37	94.22
90.95	94.38	97.72	109.45	105.95	110.79	114.89	126.72
122.69	126.53	130.60	141.54	139.21	144.76	150.04	164.28
159.92	167.94	173.86	189.70	189.53	197.09	204.49	224.77
228.33	249.29	263.02	298.85	294.17	310.03	321.31	355.67
351.57	366.96	376.99	422.29	416.80	451.51	468.25	524.60
517.65	538.89	539.46	571.94	549.71	548.97	576.00	640.11
656.81	671.92	686.82	763.11	771.61	799.53	826.25	927.02
950.67	997.64	1039.32	1167.63	1168.32	1220.98	1274.40	1450.69
1406.34	1432.50	1435.98	1646.15	1598.43	1596.53	1583.76	1810.44
1713.78	1755.75	1806.20	2120.29	1988.24	2026.31	2111.33	2392.10
2262.47	2287.23	2343.91	2695.75	2580.07	2597.08	2679.23	2979.51
2846.23	2874.44	2942.76	3310.39	3129.15	3215.51	3258.79	3571.51
3354.94	3433.27	3526.66	3860.38	3615.59	3742.75	3896.00	4518.00
4097.00	4206.00	4239.00	4824.00	4437.00	4558.00	4599.00	5375.00
4863.00	4988.00	4931.00	5442.00	5033.00	5151.00	0.00	0.00

* M1 + QUASI-MONEY (100 billion Lire, end of period)
* periodicity: QUARTERLY seasonally: UNADJUSTED starts: 60Q1
* stops: 93Q4 no. of obs: 134 last updated: 93Q2
* one line of data represents two years

132.90	136.38	142.38	151.47	152.00	155.74	161.65	175.41
174.73	179.85	187.19	202.76	205.68	210.52	216.45	208.70
206.65	210.43	215.85	235.30	234.45	241.11	248.95	271.16
270.35	276.56	285.18	306.47	306.72	314.84	323.98	348.34
346.65	357.20	367.15	393.88	395.12	403.35	414.35	443.23
444.09	460.35	471.11	513.79	508.63	529.99	550.22	601.94
606.13	630.27	650.51	716.41	720.77	763.25	791.19	881.86
895.18	928.05	939.53	1009.66	1037.29	1075.11	1130.17	1248.13
1290.80	1317.08	1356.22	1505.25	1540.40	1584.71	1634.20	1832.60
1885.20	1952.16	2022.98	2252.17	2272.46	2343.62	2427.29	2718.50
2668.30	2692.03	2716.10	3065.14	3005.41	2988.26	2984.94	3368.52
3267.19	3314.55	3409.36	3986.42	3828.29	3883.65	4015.71	4475.21
4334.15	4365.26	4487.19	5025.57	4961.01	5000.33	5120.18	5585.14
5416.11	5424.86	5545.79	6121.20	6007.11	6107.97	6162.51	6647.62
6430.61	6552.60	6736.17	7236.75	7105.68	7264.98	6634.00	7384.00

```
 6988.00 7123.00 7237.00 8020.00    7630.00 7769.00 7843.00 8781.00
 8301.00 8482.00 8455.00 9200.00    8882.00 9035.00    0.00    0.00

* MONEY SUPPLY M1 (1000 billion Lire, end of period)
* periodicity: QUARTERLY  seasonally: ADJUSTED  starts: 60Q1
* stops: 96Q4  no. of obs: 112  last updated: 96Q4
* one line of data represents two years

    0.000   0.000   0.000   0.000     0.000   0.000   0.000   0.000
    0.000   0.000   0.000   0.000     0.000   0.000   0.000   0.000
    0.000   0.000   0.000   0.000     0.000   0.000   0.000   0.000
    0.000   0.000   0.000   0.000     0.000   0.000   0.000   0.000
    0.000   0.000   0.000   0.000    19.124  19.702  20.601  21.445
   23.036  24.902  26.552  28.471    29.674  30.955  32.507  33.842
   35.413  36.640  38.178  40.160    41.865  45.113  47.471  49.910
   51.816  53.903  54.741  55.414    55.738  56.016  59.578  62.170
   66.671  68.789  71.393  74.324    78.480  82.238  86.179  89.986
   96.217 102.350 108.385 112.663   118.555 125.637 133.488 139.284
  143.071 147.902 150.811 157.039   162.407 164.910 166.124 171.981
  174.559 181.586 188.697 199.915   201.984 209.207 219.388 225.339
  229.718 235.950 242.524 253.567   262.247 267.937 276.706 280.167
  288.381 295.298 302.788 310.585   316.739 329.201 333.684 335.999
  337.108 346.668 355.165 363.078   363.798 376.372 393.300 414.300
  412.400 421.500 428.800 442.200   446.500 454.500 465.400 490.200
  488.800 497.100 498.900 497.100   506.600 512.500 528.500 532.400
  561.500 557.200 552.900 548.100   551.800 551.100 554.300 553.900
  545.800 552.200 567.100 571.900

* IMPORTS C.I.F. (billion Lire, monthly averages)
* periodicity: QUARTERLY  seasonally: ADJUSTED  starts: 60Q1
* stops: 96Q4  no. of obs: 148  last updated: 96Q4
* one line of data represents two years

   228   243   258   254     262   273   277   279
   290   302   332   337     345   392   427   416
   424   396   343   345     362   373   396   408
   431   438   455   467     487   501   511   553
   474   534   551   576     605   638   670   682
   724   777   812   805     843   791   816   850
   892   886   937  1026    1054  1301  1477  1557
  2047  2123  2484  2246    1999  1877  2166  2297
  2616  2957  3148  3445    3537  3508  3420  3611
  3434  3662  3948  4696    4751  4882  5338  6381
  6709  6681  7784  7422    7320  8950  9510  8838
 10300  9738  9740  9190    9835  9742 10551 10629
 11723 11387 12617 13698   14201 14664 13267 15329
 13870 12493 11330 11910   12661 13271 13489 14253
 13838 14912 14993 16221   17035 17778 17517 17708
 18213 17819 17766 18750   18359 18975 18909 18932
 19339 19545 18735 19259   18826 19038 18902 20769
 20669 21479 23035 24691   25738 27826 28840 29765
 27345 25849 25805 26948

* EXPORTS F.O.B. (billion Lire, monthly averages)
* periodicity: QUARTERLY  seasonally: ADJUSTED  starts: 60Q1
* stops: 96Q4  no. of obs: 148  last updated: 96Q4
* one line of data represents two years
```

181	196	190	191	205	209	223	237
236	242	244	250	253	257	269	277
283	299	328	329	357	368	384	396
397	415	429	435	455	440	445	481
496	505	535	581	601	626	631	580
640	682	718	709	746	739	815	818
884	876	878	956	854	1026	1176	1245
1418	1504	1810	1789	1836	1754	1983	2002
2200	2460	2738	2911	3116	3279	3388	3468
3365	3806	3764	4712	4613	4728	5120	5472
5493	5440	5524	5773	5744	7026	7904	7950
8510	8451	8118	8121	8626	8970	9057	10116
10677	9625	11141	11486	11976	12029	12552	13198
12652	12380	11781	11641	12259	11748	13013	12986
12762	14408	13732	14744	15405	16258	15978	16574
16930	17039	16816	17299	17558	17250	17209	17887
18445	17682	18388	18847	20490	21550	22063	24164
23995	24454	26652	26902	28925	31607	33152	33323
31992	32144	32796	32247				

APPENDIX D

COMPUTER PROGRAMS IN RATS PACKAGE

```
/* (Ia) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR TESTING
        THE MAXIMUM VAR ORDER OF THE UK MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.

UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
 UKLMO = LOG OF UK MONEY SUPPLY MO
UKLCPI = LOG OF UK CONSUMER PRICE INDEX (1990=100)
 UKTBR = UK 3-MONTH TREASURY BILL RATES
*/
CAL 1969 2 4 ;* Set CALENDAR for quarterly data beginning with 69:2
ALL 8 1996:3 ;* and ending 96:3. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\UK.DAT ;* Data set is assumed to be on drive a:\.
CLEAR UKLGDP UKLMO UKLCPI UKTBR
DATA(ORG=VAR) / UKLGDP UKLMO UKLCPI UKTBR
* The next three lines transform GDP, MO and CPI to their logs.
DOFOR I = UKLGDP UKLMO UKLCPI
  LOG I
END DOFOR
*
SMPL 1970:4 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR UKLGDP UKLMO UKLCPI UKTBR
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR UKLGDP UKLMO UKLCPI UKTBR
LAGS 1 TO 5                    <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
```

```
/* When testing a restricted VAR(l) against an unrestricted VAR(m),
l<m, the degrees of freedom are K*K*(m-l)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                    <<<<<<
# 1 TO 4
# 5 TO 8
END


Normal Completion
```

```
/* (Ib) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
        NO EC FORECASTING MODEL FOR THE UK ECONOMY

 UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
  UKLMO = LOG OF UK MONEY SUPPLY MO
 UKLCPI = LOG OF UK CONSUMER PRICE INDEX (1990=100)
  UKTBR = UK 3-MONTH TREASURY BILL RATES
*/
CAL 1969 2 4
ALL 0 1998:4
OPEN DATA A:\UK.DAT
CLEAR UKLGDP UKLMO UKLCPI UKTBR
DATA(ORG=VAR) 1969:2 1996:3 UKLGDP UKLMO UKLCPI UKTBR
*
SYSTEM(KALMAN) 1 TO 4
VAR UKLGDP UKLMO UKLCPI UKTBR
LAGS 1 TO 6
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  5.00   5.00   5.00   5.00
  2.00  40.00   0.40  24.00
300.00  24.00 300.00  15.00
  1.00   0.90   1.00   1.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1970:4 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1970:4
to 1990:4.
*/
THEIL
DO TIME = 1991:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next, the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1970:4 1996:3
THEIL
DO TIME=1996:4, 1998:4
   FORECAST 4 1
   # 1  F_UKLGDP
   # 2   F_UKLMO
   # 3  F_UKLCPI
   # 4   F_UKTBR
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1996:4 1998:4 F_UKLGDP F_UKLMO F_UKLCPI F_UKTBR
*
OPEN COPY A:\UKBV.FOR
COPY(DATES, ORG=VAR) 1996:4 1998:4 F_UKLGDP F_UKLMO F_UKLCPI F_UKTBR
END
```

```
/* (Ic) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
        EC FORECASTING MODEL FOR THE UK ECONOMY

 UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
  UKLM0 = LOG OF UK MONEY SUPPLY M0
 UKLCPI = LOG OF UK CONSUMER PRICE INDEX (1990=100)
  UKTBR = UK 3-MONTH TREASURY BILL RATES
     EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1969 2 4
ALL 0 1999:4    ;* The time span given here must be long enough.
CLEAR UKLGDP UKLM0 UKLCPI UKTBR EC
OPEN DATA A:\UK.DAT
DATA(ORG=VAR) 1969:2 1996:3 UKLGDP UKLM0 UKLCPI UKTBR
OPEN DATA A:UKRSS.DAT
DATA(ORG=VAR) 1969:2 1990:4 EC
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 1969:3 1996:3
   SET D1UKLGDP = UKLGDP(T) - UKLGDP(T-1)
   SET  D1UKLM0 =  UKLM0(T) -  UKLM0(T-1)
   SET D1UKLCPI = UKLCPI(T) - UKLCPI(T-1)
   SET  D1UKTBR =  UKTBR(T) -  UKTBR(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1UKLGDP D1UKLM0 D1UKLCPI D1UKTBR
LAGS 1 TO 5          ;* Use five lags of each variable
DET CONSTANT EC{6} ;* Include a constant and the
                     ;* error-correction term
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  1.00   1.00   1.00   0.01
 40.00  40.00  40.00   0.80
300.00   9.00 300.00  33.00
  3.00  42.00 300.00 300.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables over the period 1991:1 through
1996:3 first, so that the forecasting results can be checked against
existing data and used as guidelines for model building. At this
stage, the output from the ESTIMATE instruction is unnecessary. Use
the NOPRINT and NOFTESTS options to suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 4
THEIL(SETUP) 4 1 TIME+6
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1970:4 TIME
THEIL
DO N = TIME+1, TIME+6
```

```
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+6
    SET  TREND = T
    SET EC = UKLGDP(T) + 0.295*UKLMO(T) $
             - 1.199*UKLCPI(T) + 0.278*UKTBR(T)
 COMPUTE TIME = TIME + 6
 END DO I
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 variables concerned beyond the estimation period. Each supplementary
 card lists the equation to be used for forecasting and provides the
 dependent variable name so that the data can be extended with the
 forecasts.
 */
 COMPUTE TIME = 1996:3
 DO I = 1, 2
 THEIL(SETUP) 4 1 TIME+6
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1970:4 TIME
 THEIL
 DO N = TIME+1, TIME+6
    FORECAST 4 1
    # 1  D1UKLGDP
    # 2   D1UKLMO
    # 3  D1UKLCPI
    # 4   D1UKTBR
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+6
    SET  TREND = T
    SET UKLGDP = UKLGDP(T-1) + D1UKLGDP(T)
    SET  UKLMO =  UKLMO(T-1) +  D1UKLMO(T)
    SET UKLCPI = UKLCPI(T-1) + D1UKLCPI(T)
    SET  UKTBR =  UKTBR(T-1) +  D1UKTBR(T)
    SET EC = UKLGDP(T) + 0.295*UKLMO(T) $
             - 1.199*UKLCPI(T) + 0.278*UKTBR(T)
 COMPUTE TIME = TIME + 6
 END DO I
 *
 PRINT(DATES) 1969:2 1998:4 UKLGDP UKLMO UKLCPI UKTBR
 *
 OPEN COPY A:\UKBVEC.FOR
 COPY(DATES, ORG=VAR) 1996:4 1998:4 UKLGDP UKLMO UKLCPI UKTBR
 END
```

```
/* (IIa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR TESTING
        THE MAXIMUM VAR ORDER OF THE GERMANY MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
GERLGNP = LOG OF GERMANY REAL GDP (1985 PRICES)
 GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
GERLCPI = LOG OF GERMANY CONSUMER PRICE INDEX (1991=100)
 GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
*/
CAL 1975 3 4 ;* Set CALENDAR for quarterly data beginning with 75:3
ALL 8 1992:4 ;* and ending 92:4. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\GER.DAT ;* Data set is assumed to be on drive a:\.
CLEAR GERLGNP GERLM1 GERLCPI GERTBR
DATA(ORG=VAR) / GERLGNP GERLM1 GERLCPI GERTBR
* The next three lines transform GNP, M1 and CPI to their logs.
DOFOR I = GERLGNP GERLM1 GERLCPI
  LOG I
END DOFOR
SMPL 1977:1 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR GERLGNP GERLM1 GERLCPI GERTBR
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR GERLGNP GERLM1 GERLCPI GERTBR
LAGS 1 TO 5                        <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(1) against an unrestricted VAR(m),
1<m, the degrees of freedom are K*K*(m-1)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                        <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (IIb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
         NO EC FORECASTING MODEL FOR THE GERMAN ECONOMY

GERLGNP = LOG OF GERMANY REAL GNP (1985 PRICES)
 GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
GERLCPI = LOG OF GERMANY CONSUMER PRICE INDEX (1991=100)
 GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
*/
CAL 1975 3 4
ALL 0 1998:4
OPEN DATA A:\GER.DAT
CLEAR GERLGNP GERLM1 GERLCPI GERTBR
DATA(ORG=VAR) 1975:3 1996:3 GERLGNP GERLM1 GERLCPI GERTBR
*
SYSTEM(KALMAN) 1 TO 4
VAR GERLGNP GERLM1 GERLCPI GERTBR
LAGS 1 TO 6
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
300.00   3.00   3.00  12.00
 20.00  20.00   1.40  20.00
  1.00   1.00   1.00   1.00
 90.00   0.90   0.90  90.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables first, so that the output can be checked against existing
data and used as guidelines for model building.
*/
THEIL(SETUP) 4 1 1992:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1977:1
to 1990:4.
*/
THEIL
DO TIME=1991:1, 1992:4
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 1992:4
THEIL
```

```
DO TIME=1993:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1  F_GERLGNP
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 1996:3
THEIL
DO TIME=1996:4, 1998:4
   FORECAST 4 1
   # 1  F_GERLGNP
   #    F_GERLM1
   # 3  F_GERLCPI
   # 4   F_GERTBR
KALMAN
THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1993:1 1998:4 F_GERLGNP
PRINT(DATES) 1996:4 1998:4 F_GERLM1 F_GERLCPI F_GERTBR
*
OPEN COPY A:\GERBV1.FOR
COPY(DATES, ORG=VAR) 1993:1 1998:4 F_GERLGNP
OPEN COPY A:\GERBV2.FOR
COPY(DATES, ORG=VAR) 1996:4 1998:4 F_GERLM1 F_GERLCPI F_GERTBR
END
```

```
/* (IIc) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
         EC FORECASTING MODEL FOR THE GERMAN ECONOMY

 GERLGNP = LOG OF GERMANY REAL GNP (1985 PRICES)
  GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
 GERLCPI = LOG OF GERMANY CONSUMER PRICE INDEX (1991=100)
  GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
      EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1975 3 4
ALL 0 1999:4    ;* The time span given here must be long enough.
CLEAR GERLGNP GERLM1 GERLCPI GERTBR EC1 EC2
OPEN DATA A:\GER.DAT
DATA(ORG=VAR) 1975:3 1996:3 GERLGNP GERLM1 GERLCPI GERTBR
OPEN DATA A:GERRSS1.DAT
DATA(ORG=VAR) 1975:3 1990:4 EC1 EC2
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 1975:4 1992:4
SET D1GERLGNP = GERLGNP(T) - GERLGNP(T-1)
SMPL 1975:4 1996:3
SET  D1GERLM1 =  GERLM1(T) -  GERLM1(T-1)
SET D1GERLCPI = GERLCPI(T) - GERLCPI(T-1)
SET  D1GERTBR =  GERTBR(T) -  GERTBR(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1GERLGNP D1GERLM1 D1GERLCPI D1GERTBR
LAGS 1 TO 3                    ;* Use three lags of each variable
DET CONSTANT EC1{4} EC2{4} ;* Include a constant and the
                               ;* error-correction terms
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  1.00   0.01   1.00   1.00
  0.01   1.00   1.00   0.01
300.00   3.00 300.00  15.00
  1.00   1.00   1.00   1.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables first, so that the forecasting
results can be checked against existing data and used as guidelines
for model building. At this stage, the output from the ESTIMATE
instruction is unnecessary. Use the NOPRINT and NOFTESTS options to
suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 2
THEIL(SETUP) 4 1 TIME+4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 TIME
THEIL
```

```
 DO N = TIME+1, TIME+4
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET TREND = T
    SET   EC1 = GERLGNP(T) - 0.286*GERLM1(T) $
              + 0.086*GERLCPI(T) - 0.017*GERTBR(T)
    SET   EC2 = GERLGNP(T) - 1.590*GERLM1(T) $
              + 2.871*GERLCPI(T) - 0.024*GERTBR(T)
 COMPUTE TIME = TIME + 4
 END DO I
 *
 COMPUTE TIME = 1992:4
 DO I = 1, 4
 THEIL(SETUP) 4 1 TIME+4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 TIME
 THEIL
 DO N = TIME+1, TIME+4
    FORECAST(PRINT) 4 1
    # 1  D1GERLGNP
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET   TREND = T
    SET GERLGNP = GERLGNP(T-1) + D1GERLGNP(T)
    SET EC1 = GERLGNP(T) - 0.286*GERLM1(T) $
              + 0.086*GERLCPI(T) - 0.017*GERTBR(T)
    SET EC2 = GERLGNP(T) - 1.590*GERLM1(T) $
              + 2.871*GERLCPI(T) - 0.024*GERTBR(T)
 COMPUTE TIME = TIME + 4
 END DO I
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 variables concerned beyond the estimation period. Each supplementary
 card lists the equation to be used for forecasting and provides the
 dependent variable name so that the data can be extended with the
 forecasts.
 */
 COMPUTE TIME = 1996:3
 DO I = 1, 3
 THEIL(SETUP) 4 1 TIME+4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1977:1 TIME
```

```
THEIL
DO N = TIME+1, TIME+4
   FORECAST 4 1
   # 1  D1GERLGNP
   # 2   D1GERLM1
   # 3  D1GERLCPI
   # 4   D1GERTBR
   KALMAN
   THEIL
END DO N
TH
*
SMPL TIME+1 TIME+4
   SET   TREND = T
   SET GERLGNP = GERLGNP(T-1) + D1GERLGNP(T)
   SET  GERLM1 =  GERLM1(T-1) +  D1GERLM1(T)
   SET GERLCPI = GERLCPI(T-1) + D1GERLCPI(T)
   SET  GERTBR =  GERTBR(T-1) +  D1GERTBR(T)
   SET EC1 = GERLGNP(T) - 0.286*GERLM1(T) $
             + 0.086*GERLCPI(T) - 0.017*GERTBR(T)
   SET EC2 = GERLGNP(T) - 1.590*GERLM1(T) $
             + 2.871*GERLCPI(T) - 0.024*GERTBR(T)
COMP   TIME = TIME + 4
END D  I
*
PRINT(DATES) 1975:3 1998:4 GERLGNP GERLM1 GERLCPI GERTBR
*
OPEN COPY A:\GERBVEC.FOR
COPY(DATES, ORG=VAR) 1993:1 1998:4 GERLGNP GERLM1 GERLCPI GERTBR
END
```

```
/* (IIIa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR
          TESTING THE MAXIMUM VAR ORDER OF THE FRANCE MODEL

 In such a (modified) testing scheme, the only lines that need to be
 consecutively changed are those tagged with <<<<<<.
 FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
  FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
 FRALCPI = LOG OF FRANCE CONSUMER PRICE INDEX (1990=100)
  FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
 */
 CAL 1979 1 4 ;* Set CALENDAR for quarterly data beginning with 79:1
 ALL 8 1996:3 ;* and ending 96:3. ALLOCATE "space" of at least 4*2=8
              ;* series for the residuals.
 OPEN DATA A:\FRA.DAT ;* Data set is assumed to be on drive a:\.
 CLE  FRALGDP FRALM1 FRALCPI FRATBR
 DATA(ORG=VAR) / FRALGDP FRALM1 FRALCPI FRATBR
 * The next three lines transform GDP, M1 and CPI to their logs.
 DOFOR I = FRALGDP FRALM1 FRALCPI
   LOG I
 END DOFOR
 SMPL 1980:3 1990:4 ;* Both restricted and unrestricted VARs are
                    ;* estimated over the same sample period.
 /* The next five lines set up an 'unrestricted' VAR using 6 lags of
 each variable plus an intercept.
 */
 SYSTEM 1 TO 4
 VAR  GDP FRALM1 FRALCPI FRATBR
 LAGS  6
 DET CONSTANT
 END(SYSTEM)
 /* The next line instructs RATS to estimate the 6-lag model over the
 given sample and to save the residuals into series 1 through 4,
 respectively. At this stage, the regression output is not important;
 the options noftests and noprint cause the printing of all output to
 be suppressed.
 */
 ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
 /* Next, define a 'restricted' system using 5 lags of each variable
 and estimate the model over the same sample.
 */
 SYSTEM 1 TO 4
 VAR FRALGDP FRALM1 FRALCPI FRATBR
 LAGS 1 TO 5                          <<<<<<
 DET CONSTANT
 END(SYSTEM)
 ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
 /* When testing a restricted VAR(1) against an unrestricted VAR(m),
 1<m, the degrees of freedom are K*K*(m-1)=4*4*(6-5)=16, i.e. the
 total number of parametric restrictions imposed on the entire
 (K-dimensional) system. Also, the multiplier correction is
 K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
 unrestricted system.
 */
 RATIO(DEGREES=16, MCORR=25)                          <<<<<<
 # 1 TO 4
 # 5 TO 8
 END
```

```
/* (IIIb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
          NO EC FORECASTING MODEL FOR THE FRENCH ECONOMY

 FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
  FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
 FRALCPI = LOG OF FRANCE CONSUMER PRICE INDEX (1990=100)
  FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
*/
CAL 1979 1 4
ALL 0 1998:4
OPEN DATA A:\FRA.DAT
CLEAR FRALGDP FRALM1 FRALCPI FRATBR
DATA(ORG=VAR) 1979:1 1996:3 FRALGDP FRALM1 FRALCPI FRATBR
*
SYSTEM(KALMAN) 1 TO 4
VAR FRALGDP FRALM1 FRALCPI FRATBR
LAGS 1 TO 4
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
12.00 12.00  0.12 12.00
 6.00  6.00  6.00  0.06
 5.00  5.00  5.00  2.50
 0.01  0.01  0.60  1.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1980:1 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1980:1
to 1990:4.
*/
THEIL
DO TIME=1991:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next, the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1980:1 1996:3
THEIL
DO TIME=1996:4, 1998:4
   FORECAST 4 1
   # 1  F_FRALGDP
   # 2   F_FRALM1
   # 3  F_FRALCPI
   #4    F_FRATBR
KALMAN
THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1996:4 1998:4 F_FRALGDP F_FRALM1 F_FRALCPI F_FRATBR
*
OPEN COPY A:\FRABV.FOR
COPY(DATES, ORG=VAR) 96:4 98:4 F_FRALGDP F_FRALM1 F_FRALCPI F_FRATBR
END
```

```
/* (IIIc) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
          EC FORECASTING MODEL FOR THE FRENCH ECONOMY

 FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
  FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
 FRALCPI = LOG OF FRANCE CONSUMER PRICE INDEX (1990=100)
  FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
      EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1979 1 4
ALL 0 1999:4   ;* The time span given here must be long enough.
CLEAR FRALGDP FRALM1 FRALCPI FRATBR EC
OPEN DATA A:\FRA.DAT
DATA(ORG=VAR) 1979:1 1996:3 FRALGDP FRALM1 FRALCPI FRATBR
OPEN DATA A:FRARSS.DAT
DATA(ORG=VAR) 1979:1 1990:4 EC
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 1979:2 1996:3
   SET D1FRALGDP = FRALGDP(T) - FRALGDP(T-1)
   SET  D1FRALM1 =  FRALM1(T) -  FRALM1(T-1)
   SET D1FRALCPI = FRALCPI(T) - FRALCPI(T-1)
   SET  D1FRATBR =  FRATBR(T) -  FRATBR(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1FRALGDP D1FRALM1 D1FRALCPI D1FRATBR
LAGS 1 TO 3          ;* Use three lags of each variable
DET CONSTANT EC{4} ;* Include a constant and the
                     ;* error-correction term
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  1.00   1.00   1.00   1.00
300.00 300.00  33.00 300.00
  0.40  32.00  40.00  40.00
  0.01   0.01   0.01   1.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables over the period 1991:1 through
1996:3 first, so that the forecasting results can be checked against
existing data and used as guidelines for model building. At this
stage, the output from the ESTIMATE instruction is unnecessary. Use
the NOPRINT and NOFTESTS options to suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 6
THEIL(SETUP) 4 1 TIME+4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1980:1 TIME
THEIL
DO N = TIME+1, TIME+4
```

```
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET   TREND = T
    SET EC = FRALGDP(T) - 1.033*FRALM1(T) $
             - 0.567*FRALCPI(T) + 0.051*FRATBR(T)
 COMPUTE TIME = TIME + 4
 END DO I
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 variables concerned beyond the estimation period. Each supplementary
 card lists the equation to be used for forecasting and provides the
 dependent variable name so that the data can be extended with the
 forecasts.
 */
 COMPUTE TIME = 1996:3
 DO I = 1, 3
 THEIL(SETUP) 4 1 TIME+4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1980:1 TIME
 THEIL
 DO N = TIME+1, TIME+4
    FORECAST 4 1
    # 1  D1FRALGDP
    # 2   D1FRALM1
    # 3  D1FRALCPI
    # 4   D1FRATBR
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET   TREND = T
    SET FRALGDP = FRALGDP(T-1) + D1FRALGDP(T)
    SET  FRALM1 =  FRALM1(T-1) +  D1FRALM1(T)
    SET FRALCPI = FRALCPI(T-1) + D1FRALCPI(T)
    SET  FRATBR =  FRATBR(T-1) +  D1FRATBR(T)
    SET EC = FRALGDP(T) - 1.033*FRALM1(T) $
             - 0.567*FRALCPI(T) + 0.051*FRATBR(T)
 COMPUTE TIME = TIME + 4
 END DO I
 *
 PRINT(DATES) 1979:1 1998:4 FRALGDP FRALM1 FRALCPI FRATBR
 *
 OPEN COPY A:\FRABVEC.FOR
 COPY(DATES, ORG=VAR) 1996:4 1998:4 FRALGDP FRALM1 FRALCPI FRATBR
 END
```

```
/* (IVa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR TESTING
         THE MAXIMUM VAR ORDER OF THE ITALY MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
ITALGDP = LOG OF ITALY REAL GDP (1990 PRICES)
 ITALM1 = LOG OF ITALY MONEY SUPPLY M1
ITALCPI = LOG OF ITALY CONSUMER PRICE INDEX (1990=100)
 ITATBR = ITALY 3-MONTH TREASURY BILL RATES
*/
CAL 1977 2 4 ;* Set CALENDAR for quarterly data beginning with 77:2
ALL 8 1996:3 ;* and ending 96:3. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\ITA.DAT ;* Data set is assumed to be on drive a:\.
CLEAR ITALGDP ITALM1 ITALCPI ITATBR
DATA(ORG=VAR) / ITALGDP ITALM1 ITALCPI ITATBR
* The next three lines transform GDP, M1 and CPI to their logs.
DOFOR I = ITALGDP ITALM1 ITALCPI
  LOG I
END DOFOR
SMPL 1978:4 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR ITALGDP ITALM1 ITALCPI ITATBR
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR ITALGDP ITALM1 ITALCPI ITATBR
LAGS 1 TO 5                        <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(1) against an unrestricted VAR(m),
1<m, the degrees of freedom are K*K*(m-1)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                        <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (IVb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
         NO EC FORECASTING MODEL FOR THE ITALIAN ECONOMY

 ITALGDP = LOG OF ITALY REAL GDP (1990 PRICES)
  ITALM1 = LOG OF ITALY MONEY SUPPLY M1
 ITALCPI = LOG OF ITALY CONSUMER PRICE INDEX (1990=100)
  ITATBR = ITALY 3-MONTH TREASURY BILL RATES
*/
CAL 1977 2 4
ALL 0 1998:4
OPEN DATA A:\ITA.DAT
CLEAR ITALGDP ITALM1 ITALCPI ITATBR
DATA(ORG=VAR) 1977:2 1996:3 ITALGDP ITALM1 ITALCPI ITATBR
*
SYSTEM(KALMAN) 1 TO 4
VAR ITALGDP ITALM1 ITALCPI ITATBR
LAGS 1 TO 4
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  4.00   0.04   3.20   4.00
  6.00   6.00   0.06   6.00
  3.00   3.00 300.00   3.00
  0.01   1.00   1.00   1.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1978:2 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1978:2
to 1990:4.
*/
THEIL
DO TIME=1991:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next, the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1978:2 1996:3
THEIL
DO TIME=1996:4, 1998:4
   FORECAST 4 1
   # 1  F_ITALGDP
   # 2   F_ITALM1
   # 3  F_ITALCPI
   # 4   F_ITATBR
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1996:4 1998:4 F_ITALGDP F_ITALM1 F_ITALCPI F_ITATBR
*
OPEN COPY A:\ITABV.FOR
COPY(DATES, ORG=VAR) 96:4 98:4 F_ITALGDP F_ITALM1 F_ITALCPI F_ITATBR
END
```

```
/* (Va) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR TESTING
        THE MAXIMUM VAR ORDER OF THE LGDP MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
 UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
GERLGNP = LOG OF GERMANY REAL GNP (1985 PRICES)
FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
ITALGDP = LOG OF ITALY REAL GDP (1990 PRICES)
*/
CAL 1 71 1 4 ;* Set CALENDAR for quarterly data beginning with 71:1
ALL 8 1992:4 ;* and ending 92:4. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\GDP.DAT ;* Data set is assumed to be on drive a:\.
CLEAR UKLGDP GERLGNP FRALGDP ITALGDP
DATA(ORG=VAR) / UKLGDP GERLGNP FRALGDP ITALGDP
* The next three lines transform the selected series to their logs.
DOFOR I = UKLGDP GERLGNP FRALGDP ITALGDP
  LOG I
END DOFOR
SMPL 1972:3 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR UKLGDP GERLGNP FRALGDP ITALGDP
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR UKLGDP GERLGNP FRALGDP ITALGDP
LAGS 1 TO 5                          <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(1) against an unrestricted VAR(m),
1<m, the degrees of freedom are K*K*(m-1)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                          <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (Vb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
        NO EC FORECASTING MODEL FOR THE BIG FOUR OUTPUT GROWTH

 UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
GERLGNP = LOG OF GERMANY REAL GNP (1985 PRICES)
FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
ITALGDP = LOG OF ITALY REAL GDP (1990 PRICES)
*/
CAL 1971 1 4
ALL 0 1998:4
OPEN DATA A:\LGDP.DAT
CLEAR UKLGDP GERLGNP FRALGDP ITALGDP
DATA(ORG=VAR) 1971:1 1996:3 UKLGDP GERLGNP FRALGDP ITALGDP
*
SYSTEM(KALMAN) 1 TO 4
VAR UKLGDP GERLGNP FRALGDP ITALGDP
LAGS 1 TO 2
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
300.00  12.00 300.00   3.00
 13.00  13.00  13.00   1.43
  1.00   0.01   1.00   1.00
 24.00   3.00   3.00 300.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1992:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1971:3
to 1990:4.
*/
THEIL
DO TIME=1991:1, 1992:4
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 1992:4
```

```
 THEIL
 DO TIME=1993:1, 1996:3
    FORECAST(PRINT) 4 1
    # 1
    # 2  F_GERLGNP
    # 3
    # 4
    KALMAN
    THEIL
 END DO TIME
 THEIL(DUMP)
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 dependent variables beyond the estimation period, using explanatory
 variables that may or may not be known with certainty. Each
 supplementary card lists the equation to be used for forecasting and
 provides a name so that the forecasts can be stored for later use.
 */
 THEIL(SETUP) 4 1 1998:4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 1996:3
 THEIL
 DO TIME=1996:4, 1998:4
    FORECAST 4 1
    # 1   F_UKLGDP
    # 2  F_GERLGNP
    # 3  F_FRALGDP
    # 4  F_ITALGDP
    KALMAN
    THEIL
 END DO TIME
 THEIL(DUMP)
 *
 PRINT(DATES) 1993:1 1998:4 F_GERLGNP
 PRINT(DATES) 1996:4 1998:4 F_UKLGDP F_FRALGDP F_ITALGDP
 *
 OPEN COPY A:\GDPBV1.FOR
 COPY(DATES, ORG=VAR) 1993:1 1998:4 F_GERLGNP
 OPEN COPY A:\GDPBV2.FOR
 COPY(DATES, ORG=VAR) 1996:4 1998:4 F_UKLGDP F_FRALGDP F_ITALGDP
 END
```

```
/* (Vc) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
        EC FORECASTING MODEL FOR THE BIG FOUR OUTPUT GROWTH

  UKLGDP = LOG OF UK REAL GDP (1990 PRICES)
 GERLGNP = LOG OF GERMANY REAL GNP (1985 PRICES)
 FRALGDP = LOG OF FRANCE REAL GDP (1980 PRICES)
 ITALGDP = LOG OF ITALY REAL GDP (1990 PRICES)
      EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1971 1 4
ALL 0 1999:4    ;* The time span given here must be long enough.
CLEAR UKLGDP GERLGNP FRALGDP ITALGDP EC
OPEN DATA A:\LGDP.DAT
DATA(ORG=VAR) 1971:1 1996:3 UKLGDP GERLGNP FRALGDP ITALGDP
OPEN DATA A:LGDPRSS.DAT
DATA(ORG=VAR) 1971:1 1990:4 EC
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 71:2 1996:3
   SET  D1UKLGDP =  UKLGDP(T) -  UKLGDP(T-1)
   SET D1FRALGDP = FRALGDP(T) - FRALGDP(T-1)
   SET D1ITALGDP = ITALGDP(T) - ITALGDP(T-1)
SMPL 1971:2 1992:4
SET D1GERLGNP = GERLGNP(T) - GERLGNP(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1UKLGDP D1GERLGNP D1FRALGDP D1ITALGDP
LAG 1                ;* Use just one lag of each variable
DET CONSTANT EC(2)   ;* Include a constant and the
                     ;* error-correction term
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
300.00   3.00   3.00  21.00
  1.00   1.00   1.00   1.00
  0.01   0.01   1.00   1.00
300.00   3.00   3.00 300.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables first, so that the forecasting
results can be checked against existing data and used as guidelines
for model building. At this stage, the output from the ESTIMATE
instruction is unnecessary. Use the NOPRINT and NOFTESTS options to
suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 4
THEIL(SETUP) 4 1 TIME+2
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 TIME
THEIL
```

```
 DO N = TIME+1, TIME+2
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+2
    SET TREND = T
    SET    EC = UKLGDP(T) + 1.093*GERLGNP(T) $
              + 3.960*FRALGDP(T) - 5.360*ITALGDP(T)
 COMPUTE TIME = TIME + 2
 END DO I
 *
 COMPUTE TIME = 1992:4
 DO I = 1, 8
 THEIL(SETUP) 4 1 TIME+2
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 TIME
 THEIL
 DO N = TIME+1, TIME+2
    FORECAST(PRINT) 4 1
    # 1
    # 2  D1GERLGNP
    # 3
    # 4
    KALMAN
    THEIL
END DO N
THEIL(DUMP)
*
SMPL TIME+1 TIME+2
   SET   TREND = T
   SET GERLGNP = GERLGNP(T-1) + D1GERLGNP(T)
   SET EC = UKLGDP(T) + 1.093*GERLGNP(T) $
          + 3.960*FRALGDP(T) - 5.360*ITALGDP(T)
COMPUTE TIME = TIME + 2
END DO I
/* Next, the FORECAST instruction creates ex ante forecasts of the
variables concerned beyond the estimation period. Each supplementary
card lists the equation to be used for forecasting and provides the
dependent variable name so that the data can be extended with the
forecasts.
*/
COMPUTE TIME = 1996:3
DO I = 1, 5
THEIL(SETUP) 4 1 TIME+2
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1971:3 TIME
THEIL
DO N = TIME+1, TIME+2
   FORECAST 4 1
   # 1   D1UKLGDP
```

```
   # 2  D1GERLGNP
   # 3  D1FRALGDP
   # 4  D1ITALGDP
   KALMAN
   THEIL
END DO N
THEIL(DUMP)
*
SMPL TIME+1 TIME+2
   SET   TREND = T
   SET  UKLGDP =  UKLGDP(T-1) +  D1UKLGDP(T)
   SET GERLGNP = GERLGNP(T-1) + D1GERLGNP(T)
   SET FRALGDP = FRALGDP(T-1) + D1FRALGDP(T)
   SET ITALGDP = ITALGDP(T-1) + D1ITALGDP(T)
   SET EC = UKLGDP(T) + 1.093*GERLGNP(T) $
          + 3.960*FRALGDP(T) - 5.360*ITALGDP(T)
COMPUTE TIME = TIME + 2
END DO I
*
PRINT(DATES) 1971:1 1998:4 UKLGDP GERLGNP FRALGDP ITALGDP
*
OPEN COPY A:\LGDPBVEC.FOR
COPY(DATES, ORG=VAR) 1993:1 1998:4 UKLGDP GERLGNP FRALGDP ITALGDP
END
```

```
/* (VIa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR TESTING
        THE MAXIMUM VAR ORDER OF THE LMON MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
 UKLM0 = LOG OF UK MONEY SUPPLY M0
GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
ITALM1 = LOG OF ITALY MONEY SUPPLY M1
*/
CAL 1969 2 4 ;* Set CALENDAR for quarterly data beginning with 69:2
ALL 8 1996:4 ;* and ending 96:4. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\MON.DAT ;* Data set is assumed to be on drive a:\.
CLEAR UKLM0 GERLM1 FRALM1 ITALM1
DATA(ORG=VAR) / UKLM0 GERLM1 FRALM1 ITALM1
* The next three lines transform the selected series to their logs.
DOFOR I = UKLM0 GERLM1 FRALM1 ITALM1
  LOG I
END DOFOR
SMPL 1970:4 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR UKLM0 GERLM1 FRALM1 ITALM1
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR UKLM0 GERLM1 FRALM1 ITALM1
LAGS 1 TO 5                          <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(l) against an unrestricted VAR(m),
l<m, the degrees of freedom are K*K*(m-l)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                          <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (VIb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
         NO EC FORECASTING MODEL FOR THE BIG FOUR MONETARY GROWTH

  UKLM0 = LOG OF UK MONEY SUPPLY M0
 GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
 FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
 ITALM1 = LOG OF ITALY MONEY SUPPLY M1
 */
 CAL 1969 2 4
 ALL 0 1998:4
 OPEN DATA A:\LMON.DAT
 CLEAR UKLM0 GERLM1 FRALM1 ITALM1
 DATA(ORG=VAR) 1969:2 1996:4 UKLM0 GERLM1 FRALM1 ITALM1
 *
 SYSTEM(KALMAN) 1 TO 4
 VAR UKLM0 GERLM1 FRALM1 ITALM1
 LAGS 1 TO 4
 DET CONSTANT
 DECLARE RECT PRIORMAT(4,4)
 INPUT PRIORMAT
 300.00   3.00   3.00   3.00
   1.00   1.00   0.01   0.90
   1.00   0.60   1.00   1.00
   1.00   0.01   1.00   1.00
 SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
 END(SYSTEM)
 /* Here, prior means (MVECTOR) of the first own lags consist of a
 vector of ones, by default. In this experiment, both ex post and ex
 ante (point) forecasts are produced. The following FORECAST
 instruction prepares RATS to create ex post forecasts of the four
 variables over 1991:1-1996:4 first, so that the output can be
 checked against existing data and used as guidelines for model
 building.
 */
 THEIL(SETUP) 4 1 1996:4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1970:2 1990:4
 /* This instruction is used to suppress the printing of the OLS
 output and F-tests, and to estimate the model over the period 1970:2
 to 1990:4.
 */
 THEIL
 DO TIME=1991:1, 1996:4
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO TIME
 THEIL(DUMP)
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 dependent variables beyond the estimation period, using explanatory
 variables that may or may not be known with certainty. Each
 supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1970:2 1996:4
THEIL
DO TIME=1997:1, 1998:4
   FORECAST 4 1
   # 1   F_UKLMO
   # 2  F_GERLM1
   # 3  F_FRALM1
   # 4  F_ITALM1
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1997:1 1998:4 F_UKLMO F_GERLM1 F_FRALM1 F_ITALM1
*
OPEN COPY A:\MONBV.FOR
COPY(DATES, ORG=VAR) 97:1 98:4 F_UKLMO F_GERLM1 F_FRALM1 F_ITALM1
END
```

```
/* (VIc) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
         EC FORECASTING MODEL FOR THE BIG FOUR MONETARY GROWTH

 UKLM0 = LOG OF UK MONEY SUPPLY M0
GERLM1 = LOG OF GERMANY MONEY SUPPLY M1
FRALM1 = LOG OF FRANCE MONEY SUPPLY M1
ITALM1 = LOG OF ITALY MONEY SUPPLY M1
    EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1969 2 4
ALL 0 1998:4   ;* The time span given here must be long enough.
CLEAR UKLM0 GERLM1 FRALM1 ITALM1 EC
OPEN DATA A:\LMON.DAT
DATA(ORG=VAR) 1969:2 1996:4 UKLM0 GERLM1 FRALM1 ITALM1
OPEN DATA A:LMONRSS.DAT
DATA(ORG=VAR) 1969:2 1990:4 EC
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 1969:3 1996:4
   SET  D1UKLM1 =  UKLM0(T) -  UKLM0(T-1)
   SET D1GERLM1 = GERLM1(T) - GERLM1(T-1)
   SET D1FRALM1 = FRALM1(T) - FRALM1(T-1)
   SET D1ITALM1 = ITALM1(T) - ITALM1(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1UKLM0 D1GERLM1 D1FRALM1 D1ITALM1
LAGS 1 TO 3          ;* Use three lags of each variable
DET CONSTANT EC(4) ;* Include a constant and the
                     ;* error-correction term
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
300.00   3.00   3.00  12.00
  0.01   1.00   0.01   1.00
  1.00   0.01   1.00   1.00
  1.00   1.00   0.01   1.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables over the period 1991:1 through
1996:4 first, so that the forecasting results can be checked against
existing data and used as guidelines for model building. At this
stage, the output from the ESTIMATE instruction is unnecessary. Use
the NOPRINT and NOFTESTS options to suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 6
THEIL(SETUP) 4 1 TIME+4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1970:2 TIME
THEIL
DO N = TIME+1, TIME+4
```

```
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET   TREND = T
    SET EC = UKLMO(T) - 2.872*GERLM1(T) $
             - 4.414*FRALM1(T) + 4.323*ITALM1(T)
 COMPUTE TIME = TIME + 4
 END DO I
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 variables concerned beyond the estimation period. Each supplementary
 card lists the equation to be used for forecasting and provides the
 dependent variable name so that the data can be extended with the
 forecasts.
 */
 COMPUTE TIME = 1996:4
 DO I = 1, 2
 THEIL(SETUP) 4 1 TIME+4
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1970:2 TIME
 THEIL
 DO N = TIME+1, TIME+4
    FORECAST 4 1
    # 1   D1UKLMO
    # 2  D1GERLM1
    # 3  D1FRALM1
    # 4  D1ITALM1
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+4
    SET   TREND = T
    SET  UKLMO =  UKLMO(T-1) +  D1UKLMO(T)
    SET GERLM1 = GERLM1(T-1) + D1GERLM1(T)
    SET FRALM1 = FRALM1(T-1) + D1FRALM1(T)
    SET ITALM1 = ITALM1(T-1) + D1ITALM1(T)
    SET EC = UKLMO(T) - 2.872*GERLM1(T) $
             - 4.414*FRALM1(T) + 4.323*ITALM1(T)
 COMPUTE TIME = TIME + 4
 END DO I
 *
 PRINT(DATES) 1969:2 1998:4 UKLMO GERLM1 FRALM1 ITALM1
 *
 OPEN COPY A:\LMONBVEC.FOR
 COPY(DATES, ORG=VAR) 1991:1 1998:4 UKLMO GERLM1 FRALM1 ITALM1
 END
```

```
/* (VIIa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR
          TESTING THE MAXIMUM VAR ORDER OF THE LCPI MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
 UKLCPI = LOG OF UK CONSUMER PRICE INDEX (1990=100)
GERLCPI = LOG OF GERMANY CONSUMER PRICE INDEX (1991=100)
FRALCPI = LOG OF FRANCE CONSUMER PRICE INDEX (1990=100)
ITALCPI = LOG OF ITALY CONSUMER PRICE INDEX (1990=100)
*/
CAL 1960 1 4 ;* Set CALENDAR for quarterly data beginning with 60:1
ALL 8 1996:3 ;* and ending 96:3. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\CPI.DAT ;* Data set is assumed to be on drive a:\.
CLEAR UKLCPI GERLCPI FRALCPI ITALCPI
DATA(ORG=VAR) / UKLCPI GERLCPI FRALCPI ITALCPI
* The next three lines transform the selected series to their logs.
DOFOR I = UKLCPI GERLCPI FRALCPI ITALCPI
  LOG I
END DOFOR
SMPL 1961:3 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR UKLCPI GERLCPI FRALCPI ITALCPI
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR UKLCPI GERLCPI FRALCPI ITALCPI
LAGS 1 TO 5                        <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(l) against an unrestricted VAR(m),
l<m, the degrees of freedom are K*K*(m-l)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                        <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (VIIb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
          NO EC FORECASTING MODEL FOR THE BIG FOUR CONSUMER PRICES

 UKLCPI = LOG OF UK CONSUMER PRICE INDEX (1990=100)
GERLCPI = LOG OF GERMANY CONSUMER PRICE INDEX (1991=100)
FRALCPI = LOG OF FRANCE CONSUMER PRICE INDEX (1990=100)
ITALCPI = LOG OF ITALY CONSUMER PRICE INDEX (1990=100)
*/
CAL 1960 1 4
ALL 0 1998:4
OPEN DATA A:\LCPI.DAT
CLEAR UKLCPI GERLCPI FRALCPI ITALCPI
DATA(ORG=VAR) 1960:1 1996:3 UKLCPI GERLCPI FRALCPI ITALCPI
*
SYSTEM(KALMAN) 1 TO 4
VAR UKLCPI GERLCPI FRALCPI ITALCPI
LAGS 1 TO 4
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  9.00   1.44   4.50   9.00
300.00 300.00   3.00 300.00
300.00   3.00 300.00  90.00
 50.00  30.00  50.00  50.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1960:1 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1960:1
to 1990:4.
*/
THEIL
DO TIME=1991:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next, the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1960:1 1996:3
THEIL
DO TIME=1996:4, 1998:4
   FORECAST 4 1
   # 1   F_UKLCPI
   # 2  F_GERLCPI
   # 3  F_FRALCPI
   # 4  F_ITALCPI
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1996:4 1998:4 F_UKLCPI F_GERLCPI F_FRALCPI F_ITALCPI
*
OPEN COPY A:\CPIBV.FOR
COPY(DATES,ORG=VAR) 96:4 98:4 F_UKLCPI F_GERLCPI F_FRALCPI F_ITALCPI
END
```

```
/* (VIIIa) THE SIMS (1980) LIKELIHOOD RATIO (LR) STATISTIC FOR
           TESTING THE MAXIMUM VAR ORDER OF THE TBR MODEL

In such a (modified) testing scheme, the only lines that need to be
consecutively changed are those tagged with <<<<<<.
 UKTBR = UK 3-MONTH TREASURY BILL RATES
GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
ITATBR = ITALY 3-MONTH TREASURY BILL RATES
*/
CAL 1979 1 4 ;* Set CALENDAR for quarterly data beginning with 79:1
ALL 8 1996:3 ;* and ending 96:3. ALLOCATE "space" of at least 4*2=8
             ;* series for the residuals.
OPEN DATA A:\TBR.DAT ;* Data set is assumed to be on drive a:\.
CLEAR UKTBR GERTBR FRATBR ITATBR
DATA(ORG=VAR) / UKTBR GERTBR FRATBR ITATBR
*
SMPL 1980:3 1990:4 ;* Both restricted and unrestricted VARs are
                   ;* estimated over the same sample period.
/* The next five lines set up an 'unrestricted' VAR using 6 lags of
each variable plus an intercept.
*/
SYSTEM 1 TO 4
VAR UKTBR GERTBR FRATBR ITATBR
LAGS 1 TO 6
DET CONSTANT
END(SYSTEM)
/* The next line instructs RATS to estimate the 6-lag model over the
given sample and to save the residuals into series 1 through 4,
respectively. At this stage, the regression output is not important;
the options noftests and noprint cause the printing of all output to
be suppressed.
*/
ESTIMATE(NOPRINT, NOFTESTS) / 1 ;* Residuals into series 1 through 4
/* Next, define a 'restricted' system using 5 lags of each variable
and estimate the model over the same sample.
*/
SYSTEM 1 TO 4
VAR UKTBR GERTBR FRATBR ITATBR
LAGS 1 TO 5                        <<<<<<
DET CONSTANT
END(SYSTEM)
ESTIMATE(NOPRINT, NOFTESTS) / 5 ;* Residuals into series 5 through 8
/* When testing a restricted VAR(1) against an unrestricted VAR(m),
1<m, the degrees of freedom are K*K*(m-1)=4*4*(6-5)=16, i.e. the
total number of parametric restrictions imposed on the entire
(K-dimensional) system. Also, the multiplier correction is
K*m+1=4*6+1=25, i.e. the number of regressors per equation in the
unrestricted system.
*/
RATIO(DEGREES=16, MCORR=25)                          <<<<<<
# 1 TO 4
# 5 TO 8
END
```

```
/* (VIIIb) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
           NO EC FORECASTING MODEL FOR THE BIG FOUR TREASURY BILLS

 UKTBR = UK 3-MONTH TREASURY BILL RATES
GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
ITATBR = ITALY 3-MONTH TREASURY BILL RATES
*/
CAL 1979 1 4
ALL 0 1998:4
OPEN DATA A:\TBR.DAT
CLEAR UKTBR GERTBR FRATBR ITATBR
DATA(ORG=VAR) 1979:1 1996:3 UKTBR GERTBR FRATBR ITATBR
*
SYSTEM(KALMAN) 1 TO 4
VAR UKTBR GERTBR FRATBR ITATBR
LAGS 1 TO 2
DET CONSTANT
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
 1.00   1.00   1.00   1.00
15.00 300.00   3.00   3.00
 2.20   0.11  11.00  11.00
 6.00  60.00  30.00  60.00
SPECIFY(TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, prior means (MVECTOR) of the first own lags consist of a
vector of ones, by default. In this experiment, both ex post and ex
ante (point) forecasts are produced. The following FORECAST
instruction prepares RATS to create ex post forecasts of the four
variables over 1991:1-1996:3 first, so that the output can be
checked against existing data and used as guidelines for model
building.
*/
THEIL(SETUP) 4 1 1996:3
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1979:3 1990:4
/* This instruction is used to suppress the printing of the OLS
output and F-tests, and to estimate the model over the period 1979:3
to 1990:4.
*/
THEIL
DO TIME = 1991:1, 1996:3
   FORECAST(PRINT) 4 1
   # 1
   # 2
   # 3
   # 4
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
/* Next, the FORECAST instruction creates ex ante forecasts of the
dependent variables beyond the estimation period, using explanatory
variables that may or may not be known with certainty. Each
supplementary card lists the equation to be used for forecasting and
```

```
provides a name so that the forecasts can be stored for later use.
*/
THEIL(SETUP) 4 1 1998:4
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1979:3 1996:3
THEIL
DO TIME = 1996:4, 1998:4
   FORECAST 4 1
   # 1   F_UKTBR
   # 2  F_GERTBR
   # 3  F_FRATBR
   # 4  F_ITATBR
   KALMAN
   THEIL
END DO TIME
THEIL(DUMP)
*
PRINT(DATES) 1996:4 1998:4 F_UKTBR F_GERTBR F_FRATBR F_ITATBR
*
OPEN COPY A:\TBRBV.FOR
COPY(DATES, ORG=VAR) 96:4 98:4 F_UKTBR F_GERTBR F_FRATBR F_ITATBR
END
```

```
/* (VIIIc) THE CONSTRUCTION AND EVALUATION OF AN EMPIRICAL BVAR WITH
           EC FORECASTING MODEL FOR THE BIG FOUR TREASURY BILLS

  UKTBR = UK 3-MONTH TREASURY BILL RATES
 GERTBR = GERMANY 3-MONTH TREASURY BILL RATES
 FRATBR = FRANCE 3-MONTH TREASURY BILL RATES
 ITATBR = ITALY 3-MONTH TREASURY BILL RATES
     EC = COINTEGRATING RESIDUALS ADJUSTED FOR SHORT-RUN DYNAMICS
*/
CAL 1979 1 4
ALL 0 1999:4   ;* The time span given here must be long enough.
CLEAR UKTBR GERTBR FRATBR ITATBR EC
OPEN DATA A:\TBR.DAT
DATA(ORG=VAR) 1979:1 1996:3 UKTBR GERTBR FRATBR ITATBR
OPEN DATA A:TBRRSS2.DAT
DATA(ORG=VAR) 1979:1 1990:4 EC
SET TREND = T
/* Taking the first difference of each variable to be used in the
VAR; taking one difference loses one observation.
*/
SMPL 1979:2 1996:3
   SET  D1UKTBR =  UKTBR(T) -  UKTBR(T-1)
   SET D1GERTBR = GERTBR(T) - GERTBR(T-1)
   SET D1FRATBR = FRATBR(T) - FRATBR(T-1)
   SET D1ITATBR = ITATBR(T) - ITATBR(T-1)
* Set up a four-variable BVAR including the error-correction term.
*
SYSTEM(KALMAN) 1 TO 4
VAR D1UKTBR D1GERTBR D1FRATBR D1ITATBR
LAG 1               ;* Use just one lag of each variable
DET CONSTANT EC(2) ;* Include a constant and the
                    ;* error-correction term
DECLARE RECT PRIORMAT(4,4)
INPUT PRIORMAT
  1.00   1.00   0.01   1.00
  3.00 300.00 300.00 300.00
  3.00   3.00 300.00 300.00
  1.00   0.01   1.00   1.00
SPECIFY(MVECTOR=||0.0, 0.0, 0.0, 0.0||, $
        TYPE=GENERAL, MATRIX=PRIORMAT, TIGHT=0.01, DECAY=1.00)
END(SYSTEM)
/* Here, both ex post and ex ante (point) forecasts are produced.
The following FORECAST instruction prepares RATS to create ex post
forecasts of the four variables over the period 1991:1 through
1996:3 first, so that the forecasting results can be checked against
existing data and used as guidelines for model building. At this
stage, the output from the ESTIMATE instruction is unnecessary. Use
the NOPRINT and NOFTESTS options to suppress the display.
*/
COMPUTE TIME = 1990:4
DO I = 1, 12
THEIL(SETUP) 4 1 TIME+2
# 1 TO 4
ESTIMATE(NOPRINTS, NOFTESTS) 1979:3 TIME
THEIL
DO N = TIME+1, TIME+2
```

```
    FORECAST(PRINT) 4 1
    # 1
    # 2
    # 3
    # 4
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+2
    SET   TREND = T
    SET EC = UKTBR(T) + 107.804*GERTBR(T) $
             - 465.050*FRATBR(T) + 309.644*ITATBR(T)
 COMPUTE TIME = TIME + 2
 END DO I
 /* Next, the FORECAST instruction creates ex ante forecasts of the
 variables concerned beyond the estimation period. Each supplementary
 card lists the equation to be used for forecasting and provides the
 dependent variable name so that the data can be extended with the
 forecasts.
 */
 COMPUTE TIME = 1996:3
 DO I = 1, 5
 THEIL(SETUP) 4 1 TIME+2
 # 1 TO 4
 ESTIMATE(NOPRINTS, NOFTESTS) 1979:3 TIME
 THEIL
 DO N = TIME+1, TIME+2
    FORECAST 4 1
    # 1   D1UKTBR
    # 2  D1GERTBR
    # 3  D1FRATBR
    # 4  D1ITATBR
    KALMAN
    THEIL
 END DO N
 THEIL(DUMP)
 *
 SMPL TIME+1 TIME+2
    SET   TREND = T
    SET  UKTBR =  UKTBR(T-1) +  D1UKTBR(T)
    SET GERTBR = GERTBR(T-1) + D1GERTBR(T)
    SET FRATBR = FRATBR(T-1) + D1FRATBR(T)
    SET ITATBR = ITATBR(T-1) + D1ITATBR(T)
    SET EC = UKTBR(T) + 107.804*GERTBR(T) $
             - 465.050*FRATBR(T) + 309.644*ITATBR(T)
 COMPUTE TIME = TIME + 2
 END DO I
 *
 PRINT(DATES) 1979:1 1998:4 UKTBR GERTBR FRATBR ITATBR
 *
 OPEN COPY A:\TBRBVEC.FOR
 COPY(DATES, ORG=VAR) 1996:4 1998:4 UKTBR GERTBR FRATBR ITATBR
 END
```

APPENDIX E

DETERMINATION OF THE BEST SCALAR PRIOR HYPERPARAMETER SETTINGS

Table 6.19: Determination of the best hyperparameter setting (ν_j, w_{jk}) for Germany BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGNP	**2.86379** (3,0,0,0)	1.92519 (3,.1,0,0)	1.54057 (3,.01,.1,0)	1.33074 (3,.01,.01,.05)
	2.90135 (2,0,0,0)	1.74261 (3,.05,0,0)	1.53225 (3,.01,.05,0)	**1.32716*** (3,.01,.01,**.04**)
	2.97880 (1,0,0,0)	**1.56003** (3,**.01**,0,0)	**1.49889** (3,.01,**.01**,0)	1.32781 (3,.01,.01,.03)
LM1	0.99277 (.25,0,0,0)	0.95494 (.2,.5,0,0)	0.91804 (.2,1,.08,0)	0.97423 (.2,1,.07,.5)
	0.98875 (**.2**,0,0,0)	0.94417 (.2,.8,0,0)	**0.91792** (.2,1,**.07**,0)	0.95949 (.2,1,.07,.8)
	0.98932 (.15,0,0,0)	**0.94026** (.2,**1**,0,0)	0.91802 (.2,1,.06,0)	**0.95184*** (.2,1,.07,**1**)
LCPI	0.64803 (.1,0,0,0)	0.59359 (.01,.5,0,0)	0.34728 (.01,1,.5,0)	0.53937 (.01,1,1,.5)
	0.61086 (.05,0,0,0)	0.59348 (.01,.8,0,0)	0.34532 (.01,1,.8,0)	0.52796 (.01,1,1,.8)
	0.51049 (**.01**,0,0,0)	**0.59339** (.01,**1**,0,0)	**0.34412** (.01,1,**1**,0)	**0.51831*** (.01,1,1,**1**)
TBR	1.56406 (1,0,0,0)	2.17485 (.9,.5,0,0)	1.02033 (.9,1,.1,0)	1.00845 (.9,1,.01,.1)
	1.56346 (**.9**,0,0,0)	2.03027 (.9,.8,0,0)	0.99354 (.9,1,.05,0)	0.99853 (.9,1,.01,.05)
	1.56614 (.8,0,0,0)	**1.94532** (.9,**1**,0,0)	**0.97763** (.9,1,**.01**,0)	**0.90815*** (.9,1,.01,**.01**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.20: Determination of the best hyperparameter setting (ν_j, w_{jk}) for Germany BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGNP	0.92722 (.1,0,0,0)	0.8927158 (.01,.1,0,0)	1.06305 (.01,.01,.5,0)	1.060057 (.01,.01,1,.5)
	0.90701 (.05,0,0,0)	0.8927157 (.01,.05,0,0)	1.06227 (.01,.01,.8,0)	1.060042 (.01,.01,1,.8)
	0.89766 (**.01**,0,0,0)	**0.8927156** (.01,**.01**,0,0)	**1.06155** (.01,.01,**1**,0)	**1.060028*** (.01,.01,1,**1**)
LM1	1.43780 (.1,0,0,0)	1.474912 (.01,.1,0,0)	1.365229 (.01,.01,.5,0)	1.3343138 (.01,.01,1,.1)
	1.40937 (.05,0,0,0)	1.474908 (.01,.05,0,0)	1.365226 (.01,.01,.8,0)	1.3343114 (.01,.01,1,.05)
	1.39669 (**.01**,0,0,0)	**1.474907** (.01,**.01**,0,0)	**1.365223** (.01,.01,**1**,0)	**1.3343106*** (.01,.01,1,**.01**)
LCPI	**1.46026** (**3**,0,0,0)	1.46883 (3,.5,0,0)	1.38436 (3,1,.1,0)	1.28832 (3,1,.01,.06)
	1.46307 (2,0,0,0)	1.46384 (3,.8,0,0)	1.34707 (3,1,.05,0)	**1.28777*** (3,1,.01,**.05**)
	1.47692 (1,0,0,0)	**1.46259** (3,**1**,0,0)	**1.32589** (3,1,**.01**,0)	1.28835 (3,1,.01,.04)
TBR	0.69578 (.1,0,0,0)	0.86809 (.01,.5,0,0)	0.87647 (.01,1,.5,0)	0.98321 (.01,1,1,.5)
	0.68046 (.05,0,0,0)	0.86782 (.01,.8,0,0)	0.87644 (.01,1,.8,0)	0.98316 (.01,1,1,.8)
	0.67406 (**.01**,0,0,0)	**0.86757** (.01,**1**,0,0)	**0.87641** (.01,1,**1**,0)	**0.98312*** (.01,1,1,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.21: Determination of the best hyperparameter setting (ν_j, w_{jk}) for France BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGDP	0.95903 (.13,0,0,0)	0.91713 (.12,.5,0,0)	0.93565 (.12,1,.1,0)	0.90524 (.12,1,.01,.5)
	0.95897 (**.12**,0,0,0)	0.91651 (.12,.8,0,0)	0.92974 (.12,1,.05,0)	0.89678 (.12,1,.01,.8)
	0.95899 (.11,0,0,0)	**0.91649** (.12,**1**,0,0)	**0.92713** (.12,1,.**01**,0)	**0.89489*** (.12,1,.01,**1**)
LM1	1.18458 (.07,0,0,0)	1.18261 (.06,.5,0,0)	1.14189 (.06,1,.5,0)	1.07261 (.06,1,1,.1)
	1.18402 (**.06**,0,0,0)	1.17521 (.06,.8,0,0)	1.13915 (.06,1,.8,0)	1.07250 (.06,1,1,.05)
	1.18553 (.05,0,0,0)	**1.17339** (.06,**1**,0,0)	**1.13857** (.06,1,**1**,0)	**1.07246*** (.06,1,1,.**01**)
LCPI	0.59169 (.06,0,0,0)	0.54899 (.05,.5,0,0)	0.59418 (.05,1,.5,0)	0.45867 (.05,1,1,.6)
	0.59159 (**.05**,0,0,0)	0.53095 (.05,.8,0,0)	0.58422 (.05,1,.8,0)	**0.45865*** (.05,1,1,.**5**)
	0.59202 (.04,0,0,0)	**0.52932** (.05,**1**,0,0)	**0.57977** (.05,1,**1**,0)	0.45889 (.05,1,1,.4)
TBR	1.13103 (.1,0,0,0)	0.99233 (.01,.1,0,0)	0.991903 (.01,.01,.1,0)	0.981035 (.01,.01,.01,.7)
	1.09115 (.05,0,0,0)	0.99231 (.01,.05,0,0)	0.991900 (.01,.01,.05,0)	**0.981031*** (.01,.01,.01,**.6**)
	1.08999 (**.01**,0,0,0)	**0.99230** (.01,**.01**,0,0)	**0.991899** (.01,.01,**.01**,0)	0.98105 (.01,.01,.01,.5)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.22: Determination of the best hyperparameter setting (ν_j, w_{jk}) for France BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGDP	1.55284 (.1,0,0,0)	1.31128 (.01,.5,0,0)	1.29921 (.01,1,.5,0)	1.33078 (.01,1,1,.5)
	1.53048 (.05,0,0,0)	1.31051 (.01,.8,0,0)	1.29918 (.01,1,.8,0)	1.33077 (.01,1,1,.8)
	1.52152 **(.01**,0,0,0)	**1.30981** (.01,**1**,0,0)	**1.29916** (.01,1,**1**,0)	**1.33076*** (.01,1,1,**1**)
LM1	**1.33390** **(3**,0,0,0)	1.23842 (3,.5,0,0)	1.18137 (3,1,.12,0)	1.08817 (3,1,.11,.5)
	1.33757 (2,0,0,0)	1.23576 (3,.8,0,0)	**1.18128** (3,1,**.11**,0)	1.08408 (3,1,.11,.8)
	1.34442 (1,0,0,0)	**1.23511** (3,**1**,0,0)	1.18145 (3,1,.10,0)	**1.08309*** (3,1,.11,**1**)
LCPI	1.02107 (.5,0,0,0)	0.87544 (.4,.1,0,0)	0.76483 (.4,.01,.9,0)	0.66990 (.4,.01,.8,.5)
	1.02022 (**.4**,0,0,0)	0.86171 (.4,.05,0,0)	**0.76025** (.4,.01,**.8**,0)	0.56367 (.4,.01,.8,.8)
	1.02817 (.3,0,0,0)	**0.85747** (.4,**.01**,0,0)	0.76315 (.4,.01,.7,0)	**0.52228*** (.4,.01,.8,**1**)
TBR	3.65296 (.1,0,0,0)	2.18386 (.01,.1,0,0)	1.08928 (.01,.01,.1,0)	0.89874 (.01,.01,.01,.1)
	3.54497 (.05,0,0,0)	2.18385 (.01,.05,0,0)	1.08924 (.01,.01,.05,0)	0.898734 (.01,.01,.01,.05)
	3.49944 **(.01**,0,0,0)	**2.18384** (.01,**.01**,0,0)	**1.08923** (.01,.01,**.01**,0)	**0.898733*** (.01,.01,.01,**.01**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.23: Determination of the best hyperparameter setting (ν_j, w_{jk}) for Italy BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
LGDP	0.78721 (.05,0,0,0)	0.79987 (.04,.1,0,0)	0.75460 (.04,.01,.9,0)	0.87277 (.04,.01,.8,.5)
	0.78658 (**.04**,0,0,0)	0.79978 (.04,.05,0,0)	**0.75456** (.04,.01,**.8**,0)	0.83659 (.04,.01,.8,.8)
	0.78867 (.03,0,0,0)	**0.79974** (.04,**.01**,0,0)	0.75457 (.04,.01,.7,0)	**0.82063*** (.04,.01,.8,**1**)
LM1	0.82442 (.07,0,0,0)	0.83522 (.06,.5,0,0)	0.81418 (.06,1,.1,0)	0.84992 (.06,1,.01,.5)
	0.82438 (**.06**,0,0,0)	0.83270 (.06,.8,0,0)	0.81233 (.06,1,.05,0)	0.84580 (.06,1,.01,.8)
	0.82528 (.05,0,0,0)	**0.83219** (.06,**1**,0,0)	**0.81167** (.06,1,**.01**,0)	**0.84166*** (.06,1,.01,**1**)
LCPI	**0.57534** (**3**,0,0,0)	0.53788 (3,.1,0,0)	0.44304 (3,.01,.1,0)	0.31540 (3,.01,.01,.1)
	0.57750 (2.5,0,0,0)	0.51374 (3,.05,0,0)	0.43039 (3,.01,.05,0)	0.30233 (3,.01,.01,.05)
	0.58116 (2,0,0,0)	**0.44622** (3,**.01**,0,0)	**0.34926** (3,.01,**.01**,0)	**0.27441*** (3,.01,.01,**.01**)
TBR	1.19419 (.1,0,0,0)	1.14274 (.01,.1,0,0)	0.99983 (.01,.01,.5,0)	1.00248 (.01,.01,1,.5)
	1.18891 (.05,0,0,0)	1.142732 (.01,.05,0,0)	0.99768 (.01,.01,.8,0)	1.00062 (.01,.01,1,.8)
	1.18668 (**.01**,0,0,0)	**1.142730** (.01,**.01**,0,0)	**0.99655** (.01,.01,**1**,0)	**0.99993*** (.01,.01,1,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.24: Determination of the best hyperparameter setting (ν_j, w_{jk}) for LGDP BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKLGDP	**1.36837** (3,0,0,0)	1.09243 (3,.05,0,0)	1.08328 (3,.04,.5,0)	1.06502 (3,.04,1,.1)
	1.36885 (2,0,0,0)	**1.08268** (3,.**04**,0,0)	1.08290 (3,.04,.8,0)	1.05438 (3,.04,1,.05)
	1.37132 (1,0,0,0)	1.08786 (3,.03,0,0)	**1.08282** (3,.04,**1**,0)	**1.04295*** (3,.04,1,.**01**)
GERLGNP	0.95361 (.14,0,0,0)	0.96337 (.13,.5,0,0)	0.97996 (.13,1,.5,0)	0.9638146 (.13,1,1,.12)
	0.95276 (.**13**,0,0,0)	0.94724 (.13,.8,0,0)	0.97153 (.13,1,.8,0)	**0.9638095*** (.13,1,1,.**11**)
	0.95278 (.12,0,0,0)	**0.94331** (.13,**1**,0,0)	**0.96778** (.13,1,1,0)	0.9638098 (.13,1,1,.1)
FRALGDP	0.78251 (.1,0,0,0)	1.00607 (.01,.5,0,0)	0.84407 (.01,1,.1,0)	1.07870 (.01,1,.01,.5)
	0.77445 (.05,0,0,0)	1.00271 (.01,.8,0,0)	0.84405 (.01,1,.05,0)	1.07478 (.01,1,.01,.8)
	0.76574 (.**01**,0,0,0)	**0.99971** (.01,**1**,0,0)	**0.84404** (.01,1,.**01**,0)	**1.07193*** (.01,1,.01,**1**)
ITALGDP	**1.85606** (**3**,0,0,0)	1.83631 (3,.09,0,0)	1.73507 (3,.08,.1,0)	1.11521 (3,.08,.01,.1)
	1.85768 (2,0,0,0)	**1.83595** (3,.**08**,0,0)	1.58368 (3,.08,.05,0)	1.08803 (3,.08,.01,.05)
	1.86589 (1,0,0,0)	1.83655 (3,.07,0,0)	**1.15678** (3,.08,.**01**,0)	**1.01685*** (3,.08,.01,.**01**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.25: Determination of the best hyperparameter setting (ν_j, w_{jk}) for LGDP BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKLGDP	**1.37244** (3,0,0,0)	1.28916 (3,.1,0,0)	1.02712 (3,.01,.1,0)	0.89082 (3,.01,.01,.08)
	1.37246 (2,0,0,0)	1.16136 (3,.05,0,0)	0.98465 (3,.01,.05,0)	**0.88974*** (3,.01,.01,**.07**)
	1.37252 (1,0,0,0)	**1.05391** (3,**.01**,0,0)	**0.91960** (3,.01,**.01**,0)	0.89562 (3,.01,.01,.06)
GERLGNP	1.77483 (.1,0,0,0)	1.56676 (.01,.5,0,0)	1.367357 (.01,1,.5,0)	1.22378 (.01,1,1,.5)
	1.73601 (.05,0,0,0)	1.56661 (.01,.8,0,0)	1.367351 (.01,1,.8,0)	1.22315 (.01,1,1,.8)
	1.71735 (**.01**,0,0,0)	**1.56647** (.01,**1**,0,0)	**1.367346** (.01,1,**1**,0)	**1.22257*** (.01,1,1,**1**)
FRALGDP	1.61247 (.1,0,0,0)	1.5956028 (.01,.1,0,0)	1.5829638 (.01,.01,.1,0)	1.51656 (.01,.01,.01,.5)
	1.61017 (.05,0,0,0)	1.5956020 (.01,.05,0,0)	1.5829630 (.01,.01,.05,0)	1.51645 (.01,.01,.01,.8)
	1.60914 (**.01**,0,0,0)	**1.5956018** (.01,**.01**,0,0)	**1.5829627** (.01,.01,**.01**,0)	**1.51638*** (.01,.01,.01,**1**)
ITALGDP	**1.75977** (3,0,0,0)	1.56106 (3,.5,0,0)	1.45772 (3,1,.1,0)	1.34483 (3,1,.01,.1)
	1.76044 (2,0,0,0)	1.56028 (3,.8,0,0)	1.45365 (3,1,.05,0)	1.34315 (3,1,.01,.05)
	1.76401 (1,0,0,0)	**1.56010** (3,**1**,0,0)	**1.44585** (3,1,**.01**,0)	**1.34019*** (3,1,.01,**.01**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.26: Determination of the best hyperparameter setting (ν_j, w_{jk}) for LMON BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKLM0	**0.69115** (3,0,0,0)	0.61608 (3,.1,0,0)	0.53810 (3,.01,.1,0)	0.44847 (3,.01,.01,.1)
	0.69202 (2,0,0,0)	0.58297 (3,.05,0,0)	0.52722 (3,.01,.05,0)	0.42627 (3,.01,.01,.05)
	0.69643 (1,0,0,0)	**0.54957** (3,**.01**,0,0)	**0.48836** (3,.01,**.01**,0)	**0.40815*** (3,.01,.01,**.01**)
GERLM1	0.89145 (.1,0,0,0)	0.86187 (.01,.5,0,0)	0.848194 (.01,1,.1,0)	0.8421336 (.01,1,.01,1)
	0.88085 (.05,0,0,0)	0.86095 (.01,.8,0,0)	0.848189 (.01,1,.05,0)	**0.8421312*** (.01,1,.01,**.9**)
	0.86693 (**.01**,0,0,0)	**0.86027** (.01,**1**,0,0)	**0.848188** (.01,1,**.01**,0)	0.8421314 (.01,1,.01,.8)
FRALM1	1.70419 (.1,0,0,0)	1.42965 (.01,.5,0,0)	1.14119 (.01,1,.7,0)	1.15654 (.01,1,.6,.5)
	1.51727 (.05,0,0,0)	1.41829 (.01,.8,0,0)	**1.14118** (.01,1,.6,0)	1.14592 (.01,1,.6,.8)
	1.38944 (**.01**,0,0,0)	**1.41023** (.01,**1**,0,0)	1.14125 (.01,1,.5,0)	**1.14089*** (.01,1,.6,**1**)
ITALM1	1.00104 (.1,0,0,0)	0.85919 (.01,.5,0,0)	0.90250 (.01,1,.1,0)	0.90355 (.01,1,.01,.5)
	0.95457 (.05,0,0,0)	0.85895 (.01,.8,0,0)	0.90243 (.01,1,.05,0)	0.90044 (.01,1,.01,.8)
	0.88540 (**.01**,0,0,0)	**0.85876** (.01,**1**,0,0)	**0.90241** (.01,1,**.01**,0)	**0.89803*** (.01,1,.01,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.27: Determination of the best hyperparameter setting (ν_j, w_{jk}) for LMON BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKLM0	**2.57628** (3,0,0,0)	2.06291 (3,.1,0,0)	0.75129 (3,.01,.1,0)	0.68725 (3,.01,.01,.05)
	2.58030 (2,0,0,0)	1.42145 (3,.05,0,0)	0.72152 (3,.01,.05,0)	**0.68656*** (3,.01,.01,**.04**)
	2.60155 (1,0,0,0)	**0.82206** (3,**.01**,0,0)	**0.70881** (3,.01,**.01**,0)	0.68841 (3,.01,.01,.03)
GERLM1	0.72328 (.1,0,0,0)	0.7010979 (.01,.1,0,0)	0.6932006 (.01,.01,.1,0)	0.6934116 (.01,.01,.01,.5)
	0.71169 (.05,0,0,0)	0.7010974 (.01,.05,0,0)	0.69320054 (.01,.01,.05,0)	0.6934113 (.01,.01,.01,.8)
	0.70770 (**.01**,0,0,0)	**0.7010972** (.01,**.01**,0,0)	**0.69320051** (.01,.01,**.01**,0)	**0.6934110*** (.01,.01,.01,**1**)
FRALM1	1.76415 (.1,0,0,0)	1.81823 (.01,.5,0,0)	1.14765 (.01,1,.1,0)	1.28943 (.01,1,.01,.5)
	1.73302 (.05,0,0,0)	1.81812 (.01,.8,0,0)	1.147631 (.01,1,.05,0)	1.28923 (.01,1,.01,.8)
	1.72086 (**.01**,0,0,0)	**1.81803** (.01,**1**,0,0)	**1.147626** (.01,1,**.01**,0)	**1.28905*** (.01,1,.01,**1**)
ITALM1	1.11951 (.1,0,0,0)	1.11379 (.01,.5,0,0)	1.29290 (.01,1,.5,0)	0.867562 (.01,1,1,.1)
	1.10765 (.05,0,0,0)	1.11373 (.01,.8,0,0)	1.29281 (.01,1,.8,0)	0.867557 (.01,1,1,.05)
	1.10372 (**.01**,0,0,0)	**1.11369** (.01,**1**,0,0)	**1.29272** (.01,1,**1**,0)	**0.867555*** (.01,1,1,**.01**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.28: Determination of the best hyperparameter setting (ν_j, w_{jk}) for LCPI BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKLCPI	0.69100 (.1,0,0,0)	0.70990 (.09,.17,0,0)	0.72808 (.09,.16,.6,0)	0.80795 (.09,.16,.5,.5)
	0.69084 (**.09**,0,0,0)	**0.70984** (.09,**.16**,0,0)	**0.72807** (.09,.16,**.5**,0)	0.78264 (.09,.16,.5,.8)
	0.69090 (.08,0,0,0)	0.70996 (.09,.15,0,0)	0.72841 (.09,.16,.4,0)	**0.77230*** (.09,.16,.5,**1**)
GERLCPI	**0.57594** (3,0,0,0)	0.58510 (3,.5,0,0)	0.57728 (3,1,.1,0)	0.52348 (3,1,.01,.5)
	0.57598 (2,0,0,0)	0.57978 (3,.8,0,0)	0.57452 (3,1,.05,0)	0.523459 (3,1,.01,.8)
	0.57671 (1,0,0,0)	**0.57844** (3,**1**,0,0)	**0.52346** (3,1,**.01**,0)	**0.523457*** (3,1,.01,**1**)
FRALCPI	**0.76242** (3,0,0,0)	0.76331 (3,.5,0,0)	0.69796 (3,1,.1,0)	0.58562 (3,1,.01,.4)
	0.76245 (2,0,0,0)	0.76274 (3,.8,0,0)	0.65591 (3,1,.05,0)	**0.58498*** (3,1,.01,**.3**)
	0.76260 (1,0,0,0)	**0.76262** (3,**1**,0,0)	**0.58755** (3,1,**.01**,0)	0.58509 (3,1,.01,.2)
ITALCPI	0.26857 (.6,0,0,0)	0.28823 (.5,.5,0,0)	0.25743 (.5,1,.7,0)	0.26221 (.5,1,.6,.5)
	0.26844 (**.5**,0,0,0)	0.28131 (.5,.8,0,0)	**0.25713** (.5,1,**.6**,0)	0.25876 (.5,1,.6,.8)
	0.26852 (.4,0,0,0)	**0.27798** (.5,**1**,0,0)	0.25732 (.5,1,.5,0)	**0.25803*** (.5,1,.6,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.29: Determination of the best hyperparameter setting (ν_j, w_{jk}) for TBR BVAR forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKTBR	1.07794 (.1,0,0,0)	1.12446 (.01,.5,0,0)	1.09650 (.01,1,.5,0)	0.98371 (.01,1,1,.5)
	0.98548 (.05,0,0,0)	1.12322 (.01,.8,0,0)	1.09645 (.01,1,.8,0)	0.98345 (.01,1,1,.8)
	0.95928 (.01,0,0,0)	**1.12211** (.01,1,0,0)	**1.09640** (.01,1,1,0)	**0.98331*** (.01,1,1,1)
GERTBR	**1.34050** (3,0,0,0)	1.29856 (3,.06,0,0)	1.11137 (3,.05,.1,0)	0.89820 (3,.05,.01,.1)
	1.34117 (2,0,0,0)	**1.29659** (3,**.05**,0,0)	1.03664 (3,.05,.05,0)	0.86150 (3,.05,.01,.05)
	1.34488 (1,0,0,0)	1.29736 (3,.04,0,0)	**0.97746** (3,.05,**.01**,0)	**0.83390*** (3,.05,.01,**.01**)
FRATBR	0.90057 (.12,0,0,0)	0.90077 (.11,.25,0,0)	0.93730 (.11,.2,.1,0)	1.01695 (.11,.2,.01,.5)
	0.90034 (**.11**,0,0,0)	**0.90071** (.11,**.2**,0,0)	0.93648 (.11,.2,.05,0)	1.00159 (.11,.2,.01,.8)
	0.90085 (.10,0,0,0)	0.90072 (.11,.15,0,0)	**0.93621** (.11,.2,**.01**,0)	**0.99298*** (.11,.2,.01,**1**)
ITATBR	0.94160 (.7,0,0,0)	0.90461 (.6,.1,0,0)	0.92625 (.6,.01,.5,0)	0.90811 (.6,.01,1,.6)
	0.94158 (**.6**,0,0,0)	0.89634 (.6,.05,0,0)	0.91715 (.6,.01,**.8**,0)	**0.90781*** (.6,.01,1,.5)
	0.94168 (.5,0,0,0)	**0.89236** (.6,**.01**,0,0)	**0.91351** (.6,.01,1,0)	0.90801 (.6,.01,1,.4)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

Table 6.30: Determination of the best hyperparameter setting (ν_j, w_{jk}) for TBR BVAR-EC forecasting model

Dep Var	Theil U Statistics for One-Step-Ahead Forecasts			
	$(\nu_j, 0, 0, 0)$	$(\nu_j, w_{j1}, 0, 0)$	$(\nu_j, w_{j1}, w_{j2}, 0)$	$(\nu_j, w_{j1}, w_{j2}, w_{j3})$
UKTBR	1.59373 (.1,0,0,0)	1.60929 (.01,.5,0,0)	1.447984 (.01,1,.1,0)	1.48018 (.01,1,.01,.5)
	1.59369 (.05,0,0,0)	1.60926 (.01,.8,0,0)	1.447981 (.01,1,.05,0)	1.48009 (.01,1,.01,.8)
	1.59368 **(.01**,0,0,0)	**1.60923** (.01,**1**,0,0)	**1.447980** (.01,1,.**01**,0)	**1.48001*** (.01,1,.01,**1**)
GERTBR	**0.85311** (3,0,0,0)	0.85255 (3,.1,0,0)	0.85070 (3,.01,.5,0)	0.85196 (3,.01,1,.5)
	0.85315 (2,0,0,0)	0.85171 (3,.05,0,0)	0.85054 (3,.01,.8,0)	0.85107 (3,.01,1,.8)
	0.85334 (1,0,0,0)	**0.85043** (3,**.01**,0,0)	**0.85050** (3,.01,**1**,0)	**0.85087*** (3,.01,1,**1**)
FRATBR	**1.29980** (3,0,0,0)	1.29938 (3,.1,0,0)	1.19664 (3,.01,.1,0)	1.09528 (3,.01,.01,.5)
	1.30038 (2,0,0,0)	1.29875 (3,.05,0,0)	1.19498 (3,.01,.05,0)	1.09374 (3,.01,.01,.8)
	1.30346 (1,0,0,0)	**1.29780** (3,**.01**,0,0)	**1.19274** (3,.01,**.01**,0)	1.09338* (3,.01,.01,**1**)
ITATBR	0.73991 (.1,0,0,0)	0.74207 (.01,.5,0,0)	0.7267013 (.01,1,.1,0)	0.80032 (.01,1,.01,.5)
	0.73730 (.05,0,0,0)	0.74200 (.01,.8,0,0)	0.7267007 (.01,1,.05,0)	0.80015 (.01,1,.01,.8)
	0.73615 (**.01**,0,0,0)	**0.74199** (.01,**1**,0,0)	**0.7267005** (.01,1,.**01**,0)	**0.79999*** (.01,1,.01,**1**)

Note: * indicates the minimum value in Theil U statistics of each equation associated with the chosen hyperparameters during the ex post forecast period.

APPENDIX F

FORECASTING RESULTS

This appendix provides graphic displays of the BVAR and BVAR-EC quarterly forecasts made over 1991Q1-1998Q4 for Germany, France and Italy.

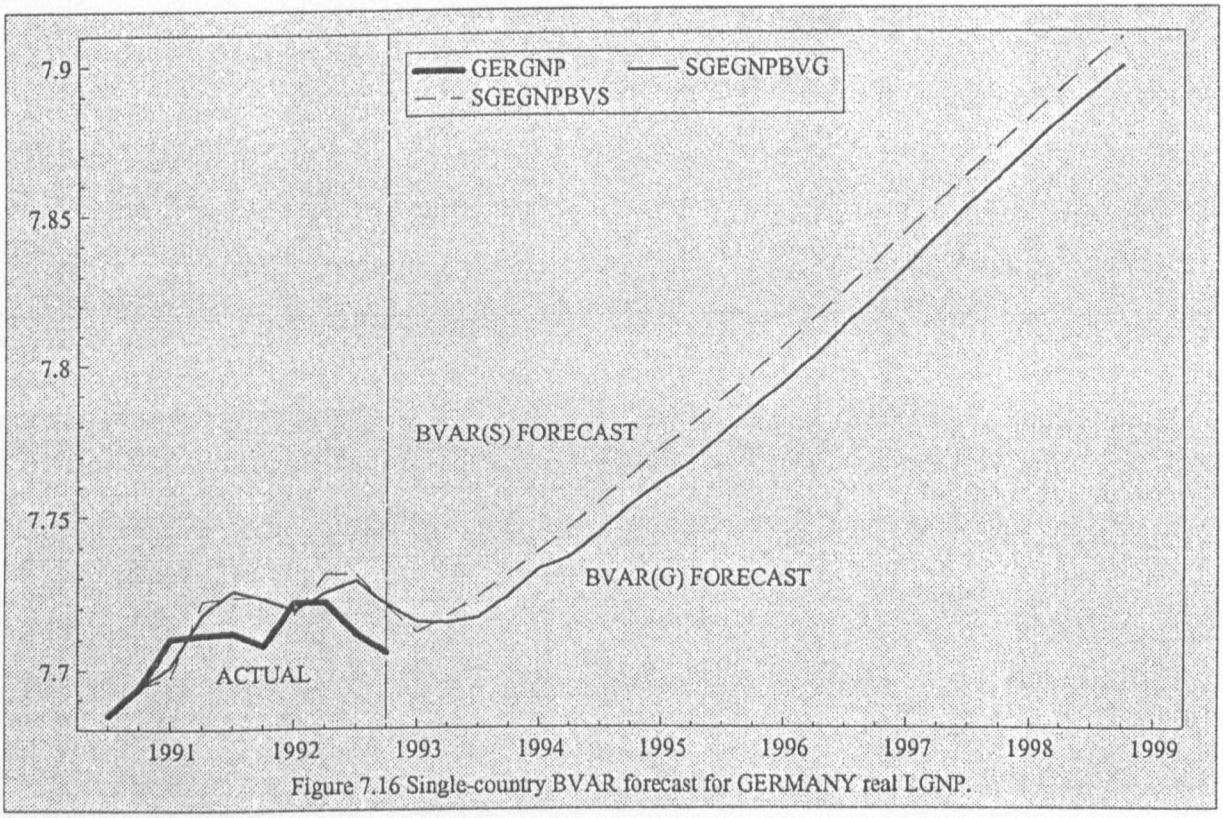

Figure 7.16 Single-country BVAR forecast for GERMANY real LGNP.

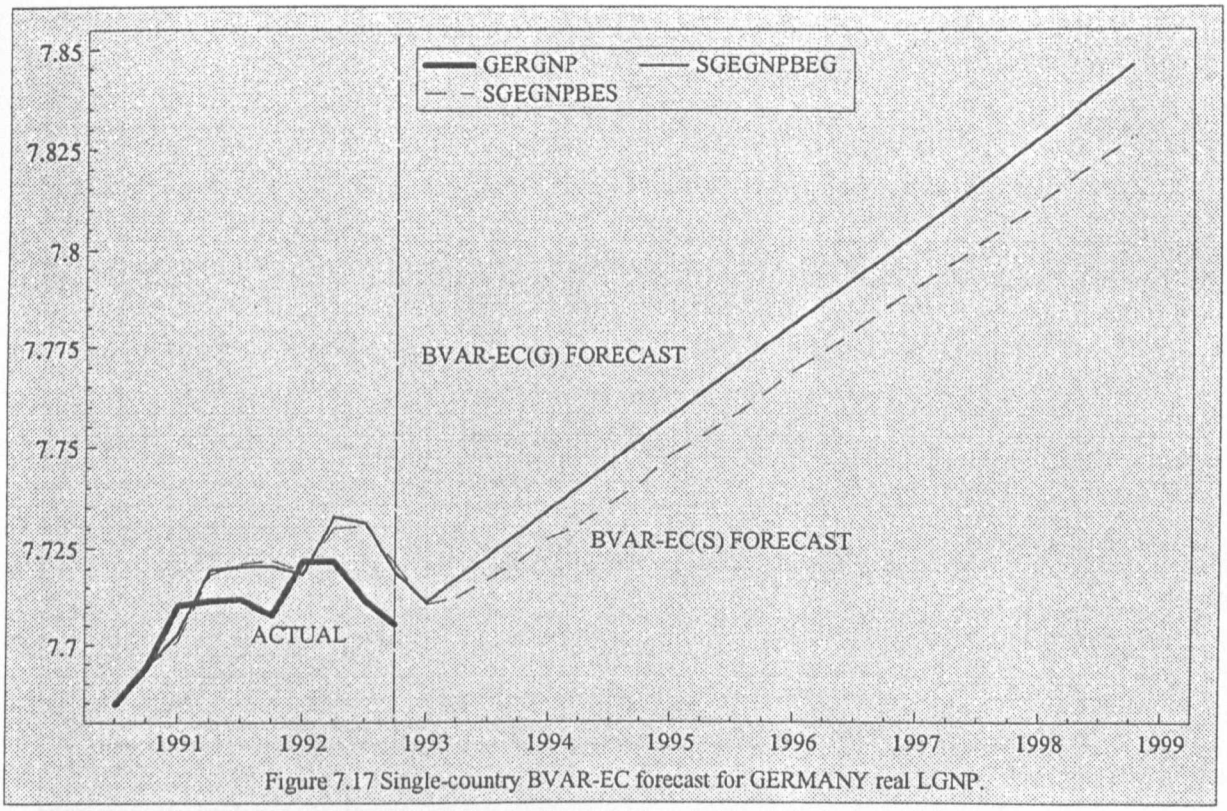

Figure 7.17 Single-country BVAR-EC forecast for GERMANY real LGNP.

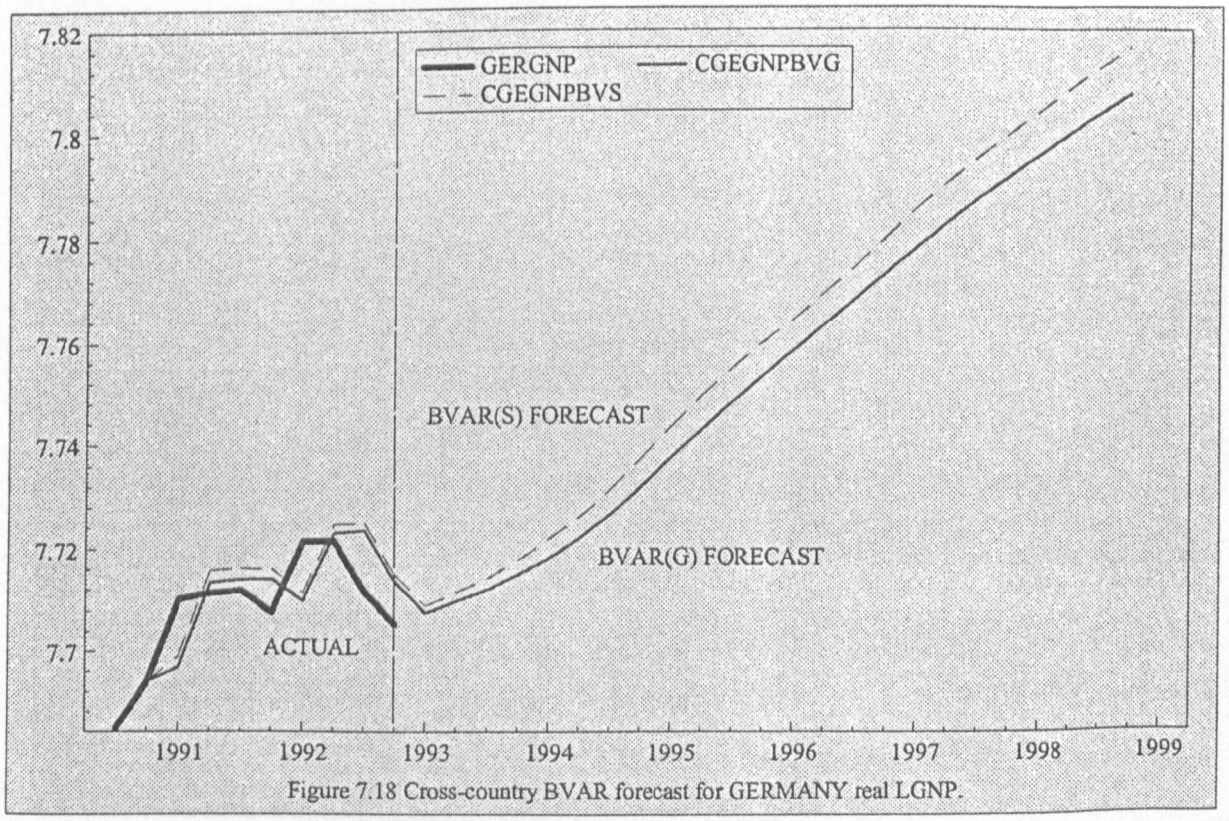

Figure 7.18 Cross-country BVAR forecast for GERMANY real LGNP.

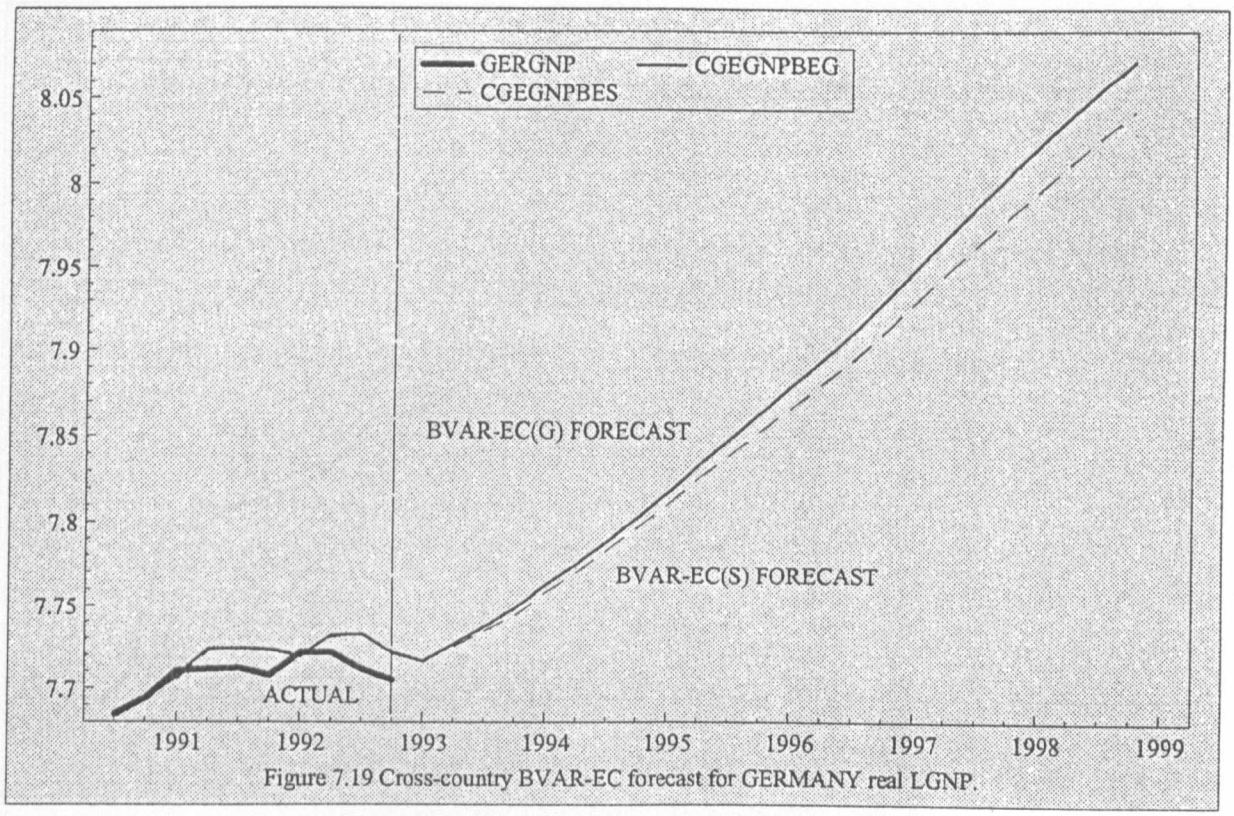

Figure 7.19 Cross-country BVAR-EC forecast for GERMANY real LGNP.

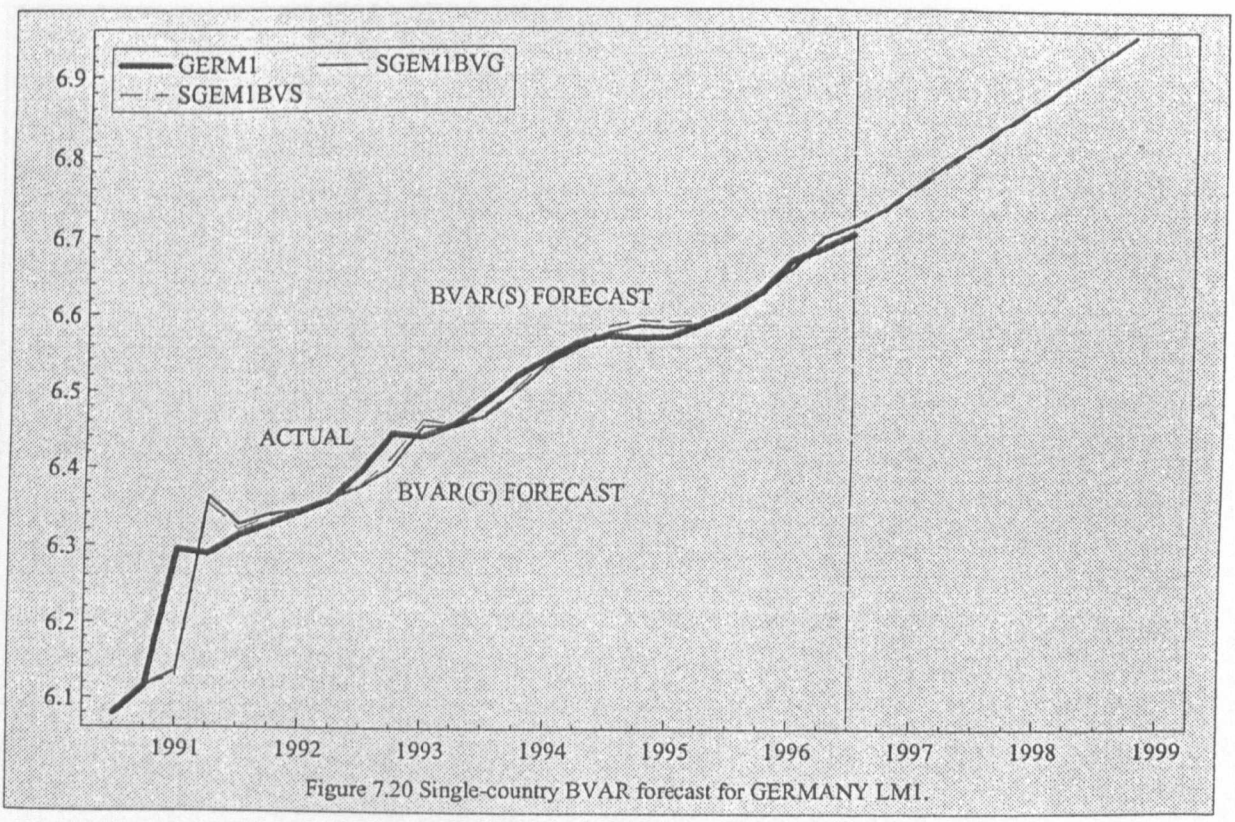

Figure 7.20 Single-country BVAR forecast for GERMANY LM1.

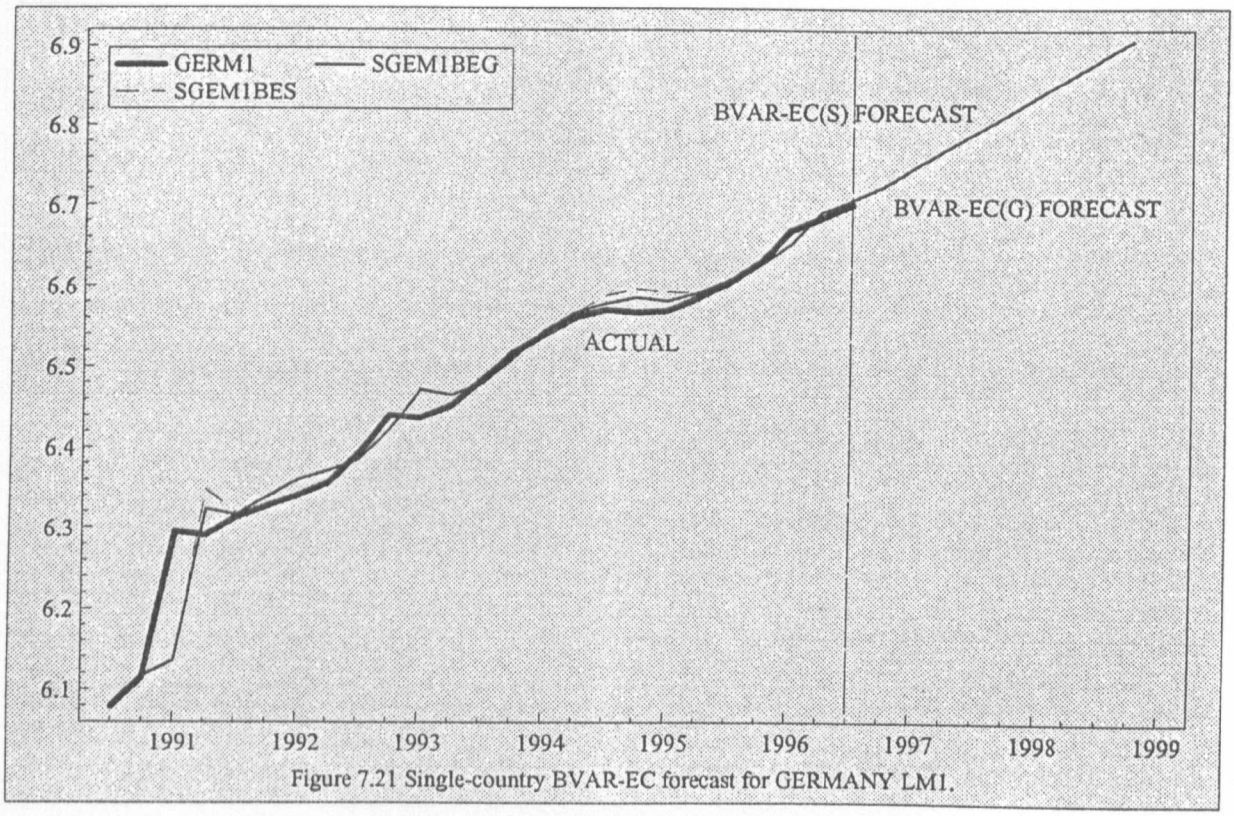

Figure 7.21 Single-country BVAR-EC forecast for GERMANY LM1.

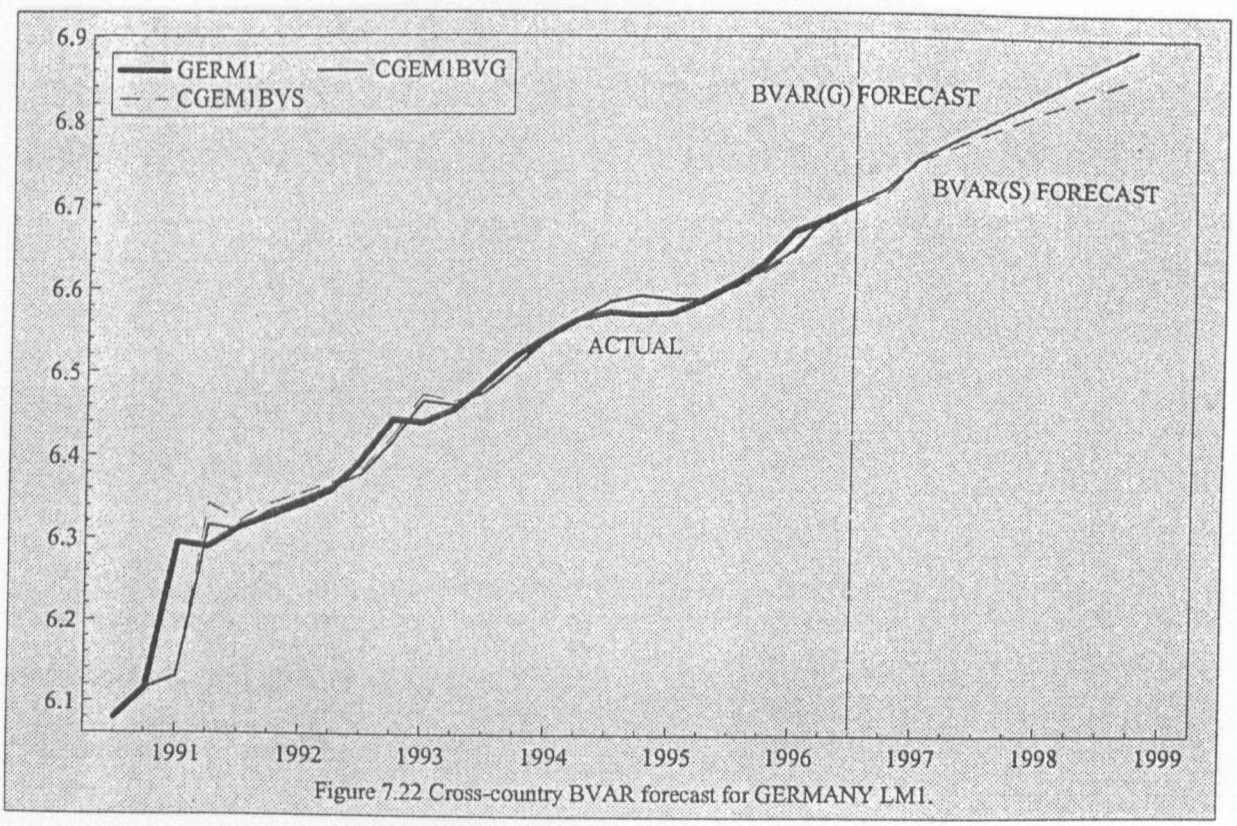

Figure 7.22 Cross-country BVAR forecast for GERMANY LM1.

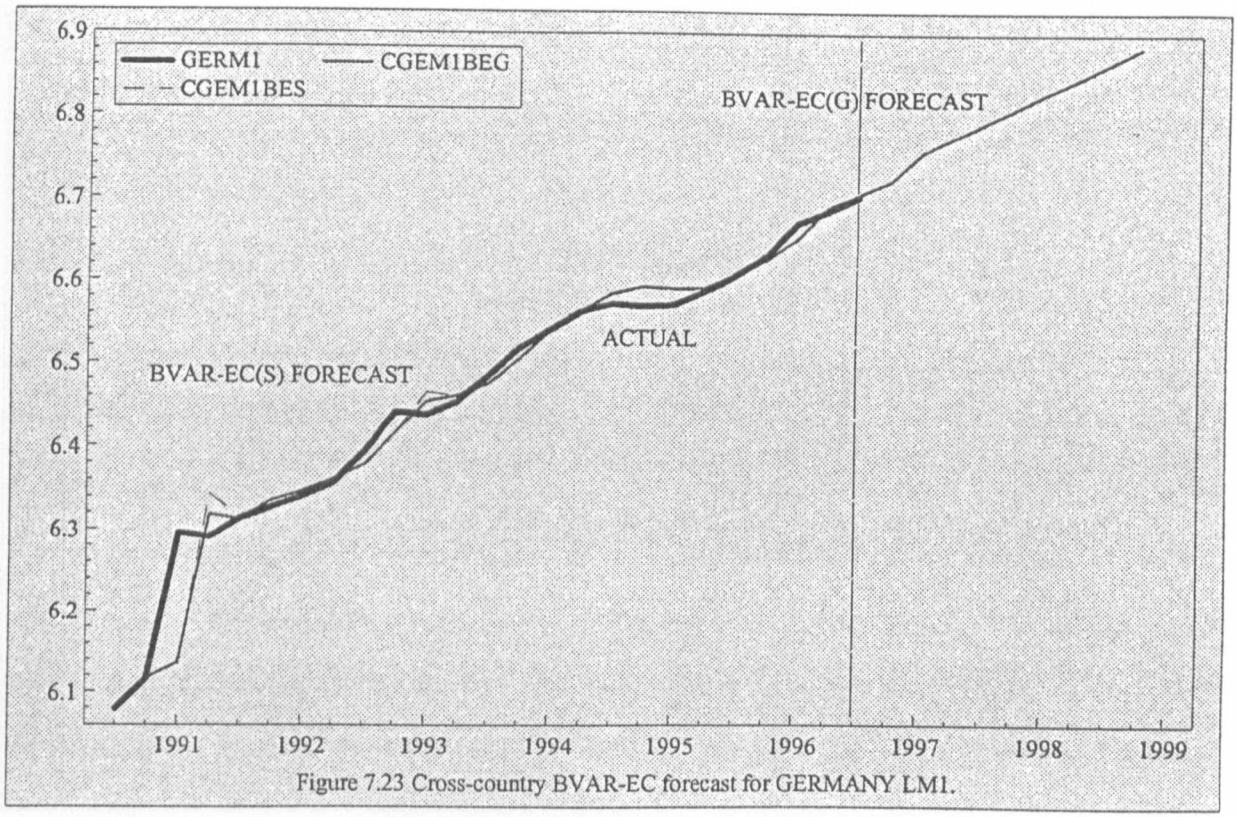

Figure 7.23 Cross-country BVAR-EC forecast for GERMANY LM1.

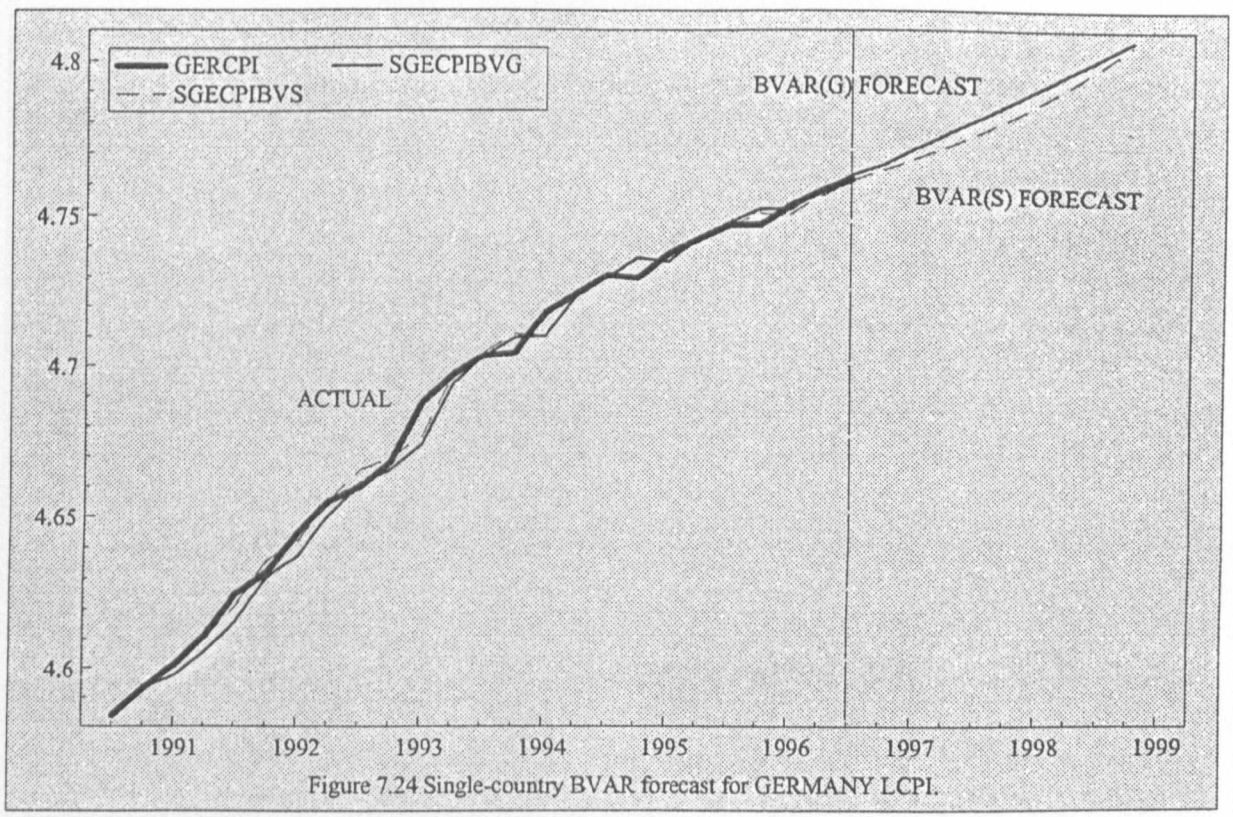

Figure 7.24 Single-country BVAR forecast for GERMANY LCPI.

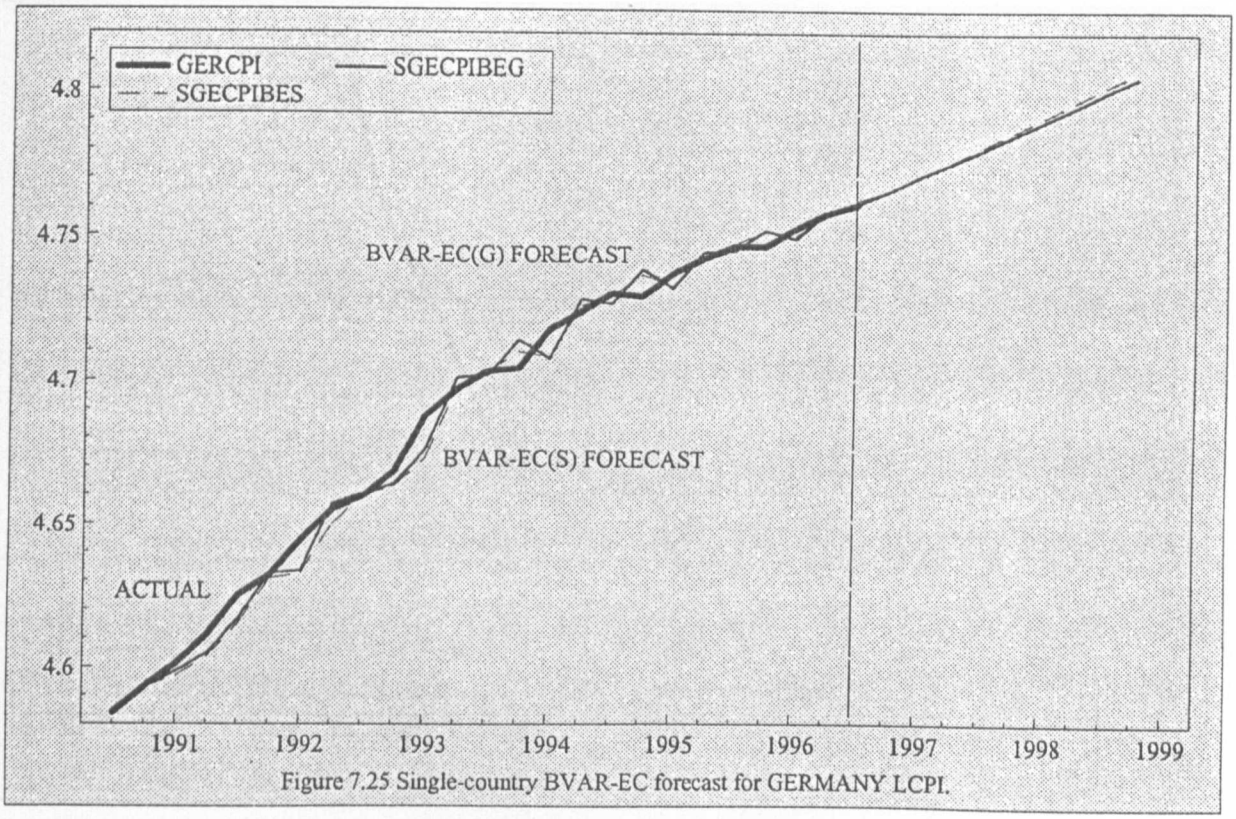

Figure 7.25 Single-country BVAR-EC forecast for GERMANY LCPI.

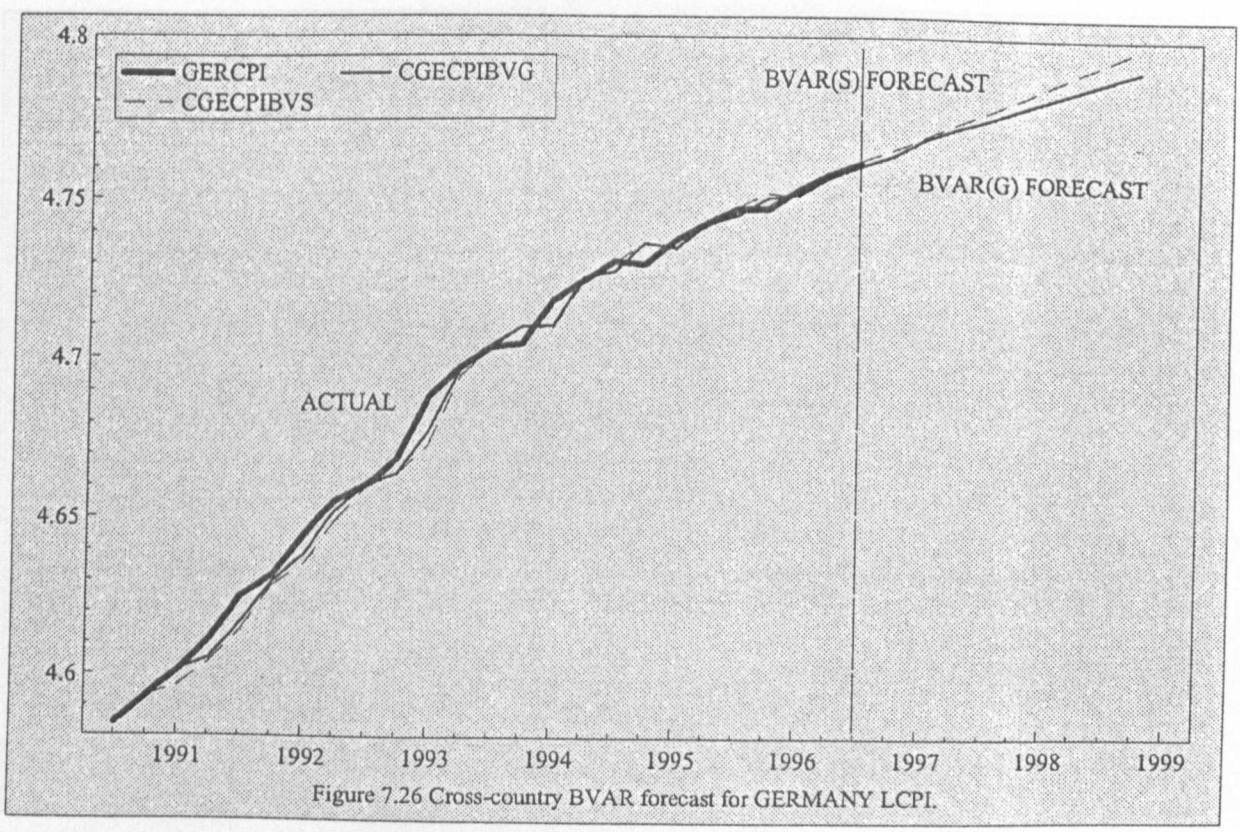

Figure 7.26 Cross-country BVAR forecast for GERMANY LCPI.

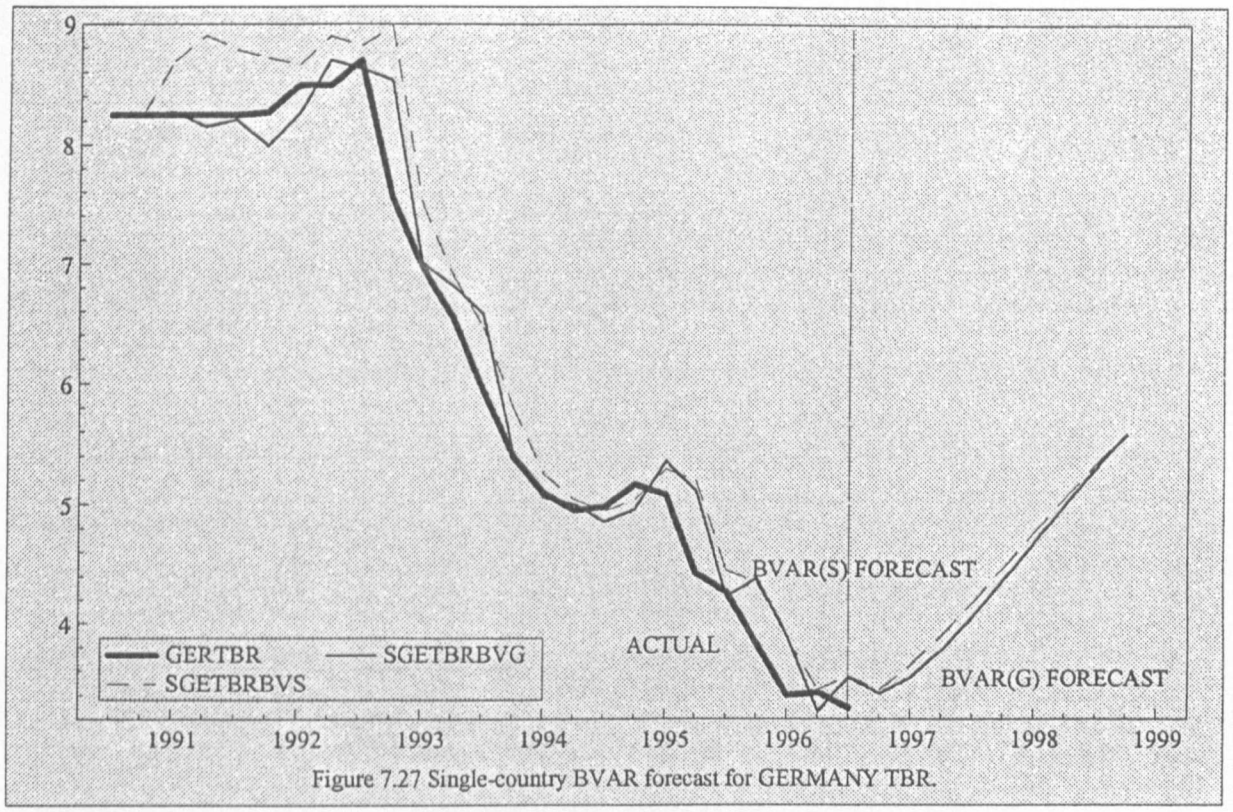

Figure 7.27 Single-country BVAR forecast for GERMANY TBR.

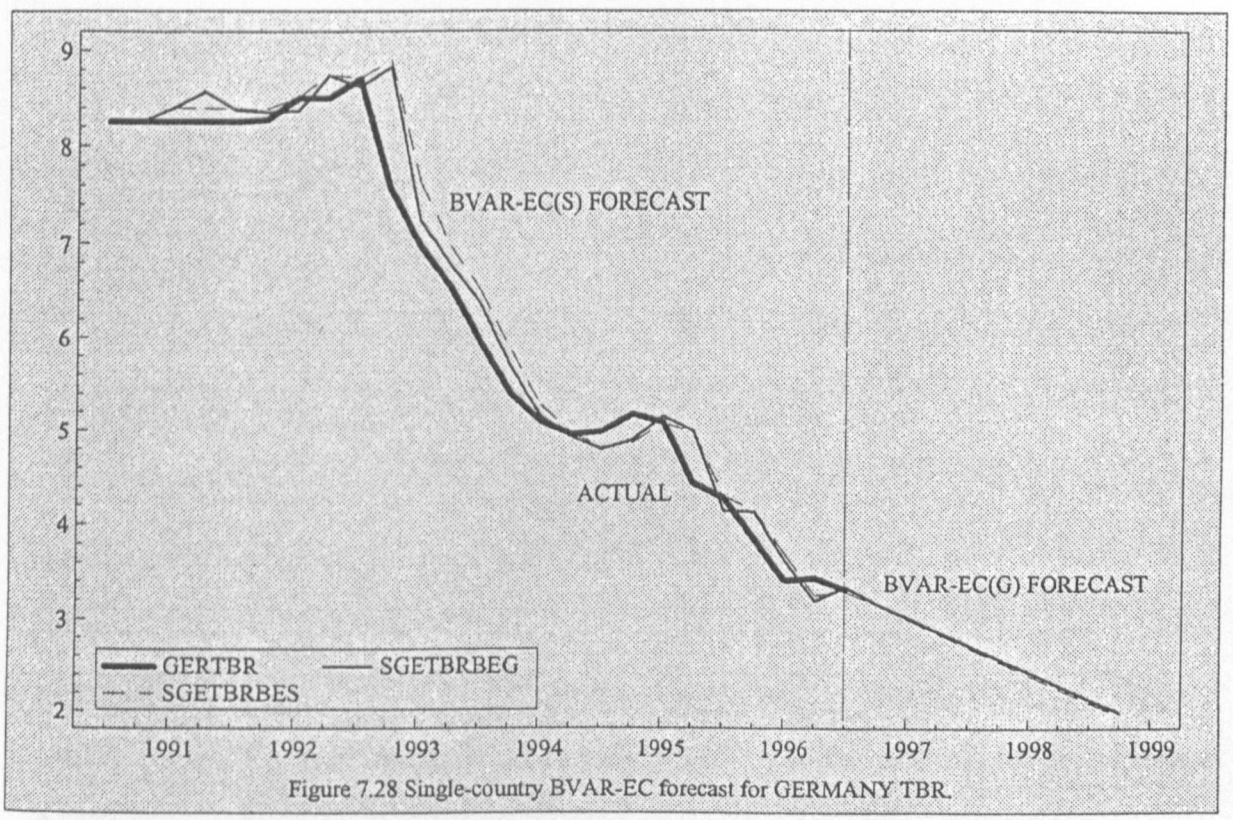

Figure 7.28 Single-country BVAR-EC forecast for GERMANY TBR.

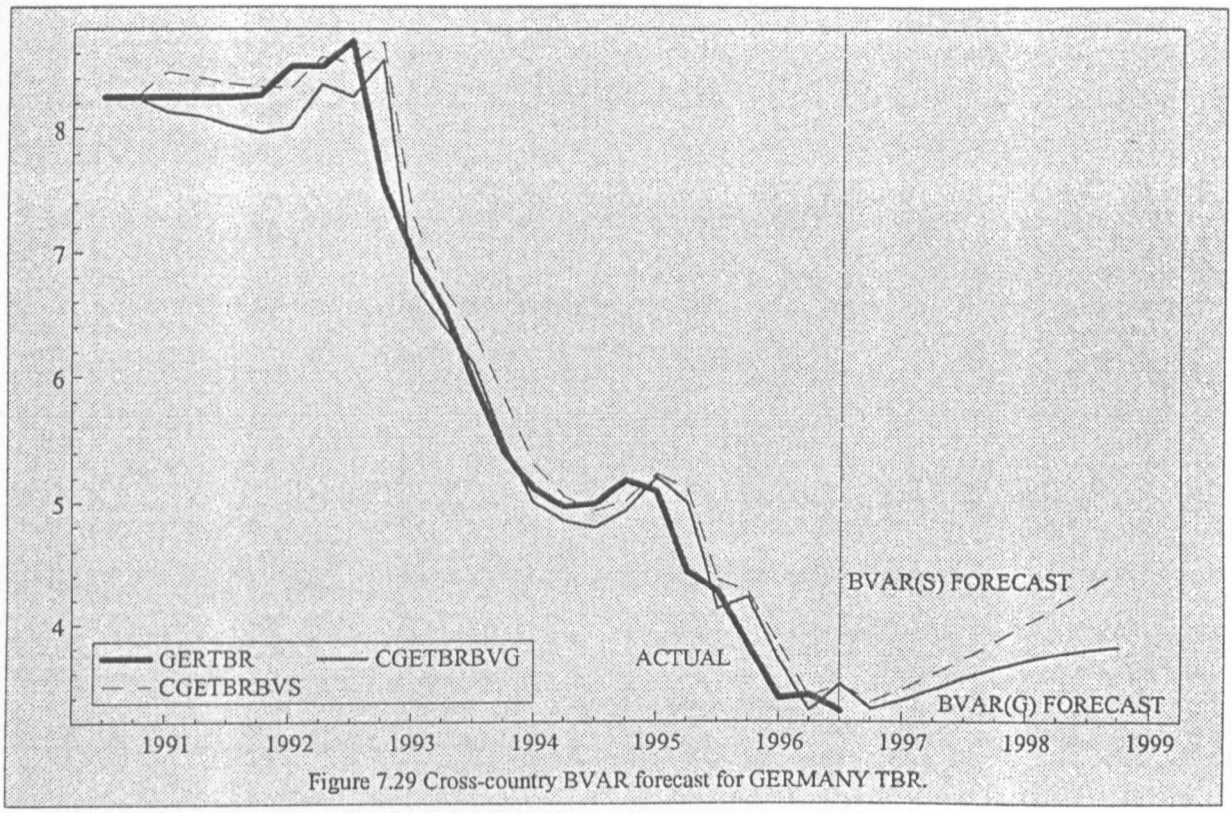

Figure 7.29 Cross-country BVAR forecast for GERMANY TBR.

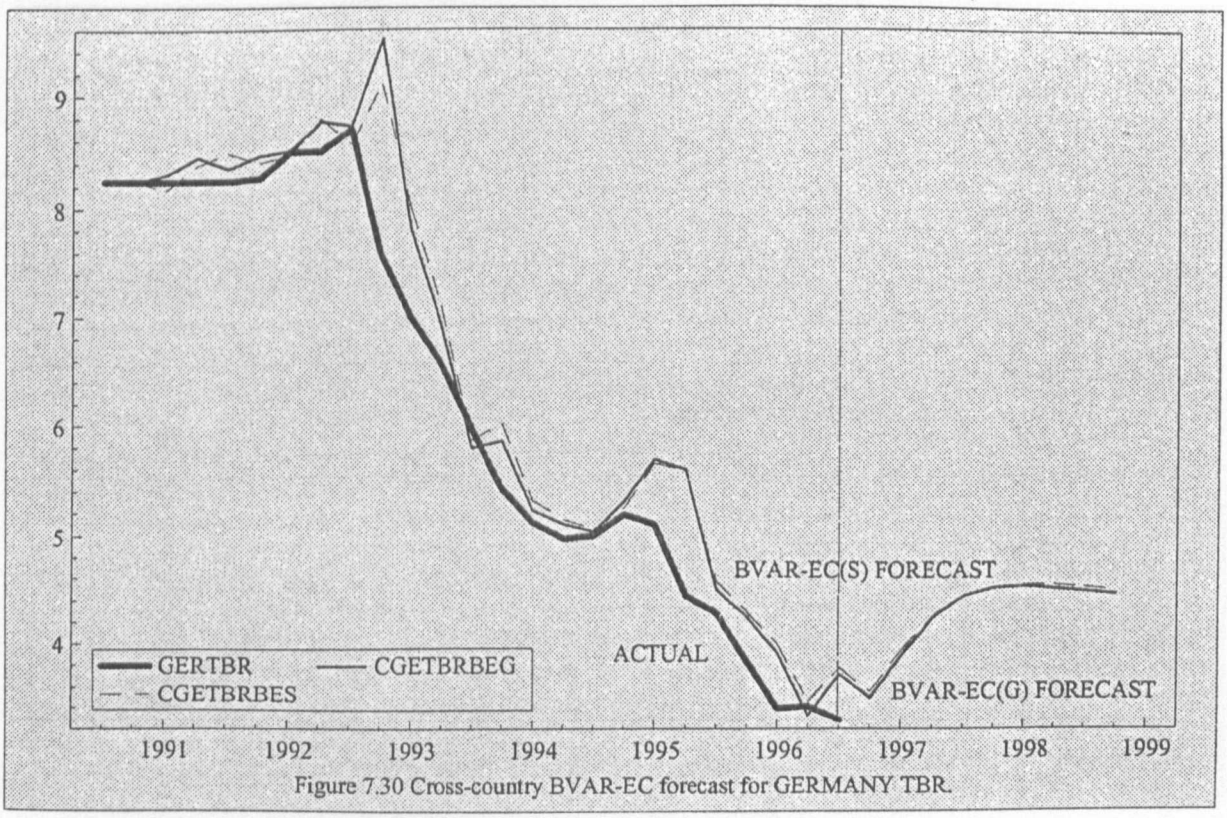

Figure 7.30 Cross-country BVAR-EC forecast for GERMANY TBR.

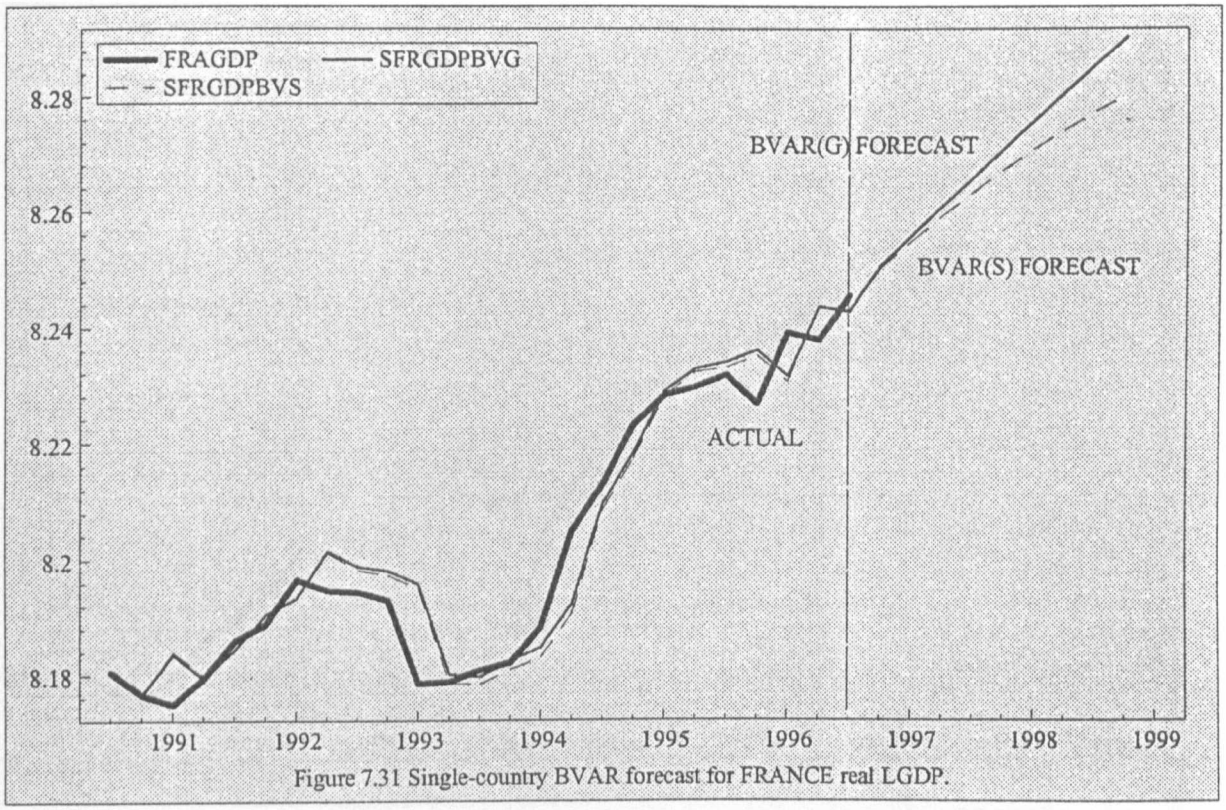

Figure 7.31 Single-country BVAR forecast for FRANCE real LGDP.

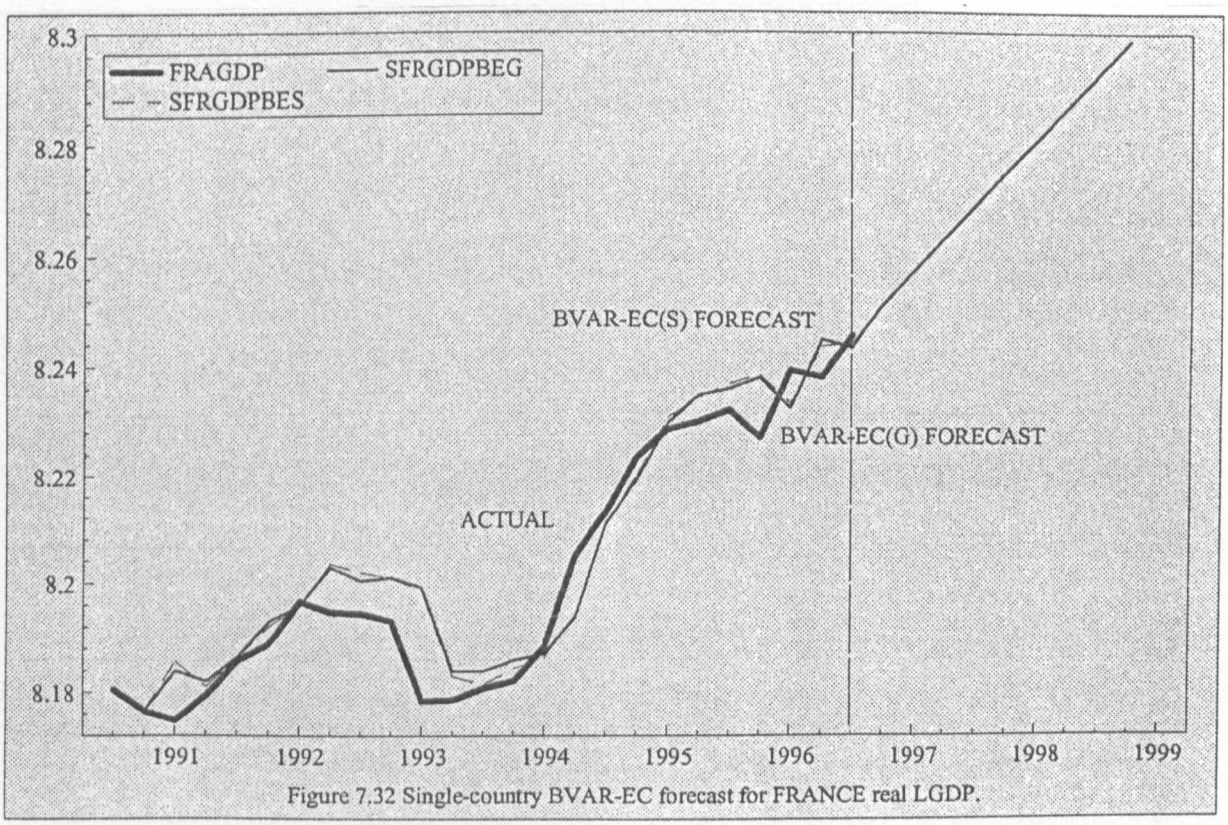

Figure 7.32 Single-country BVAR-EC forecast for FRANCE real LGDP.

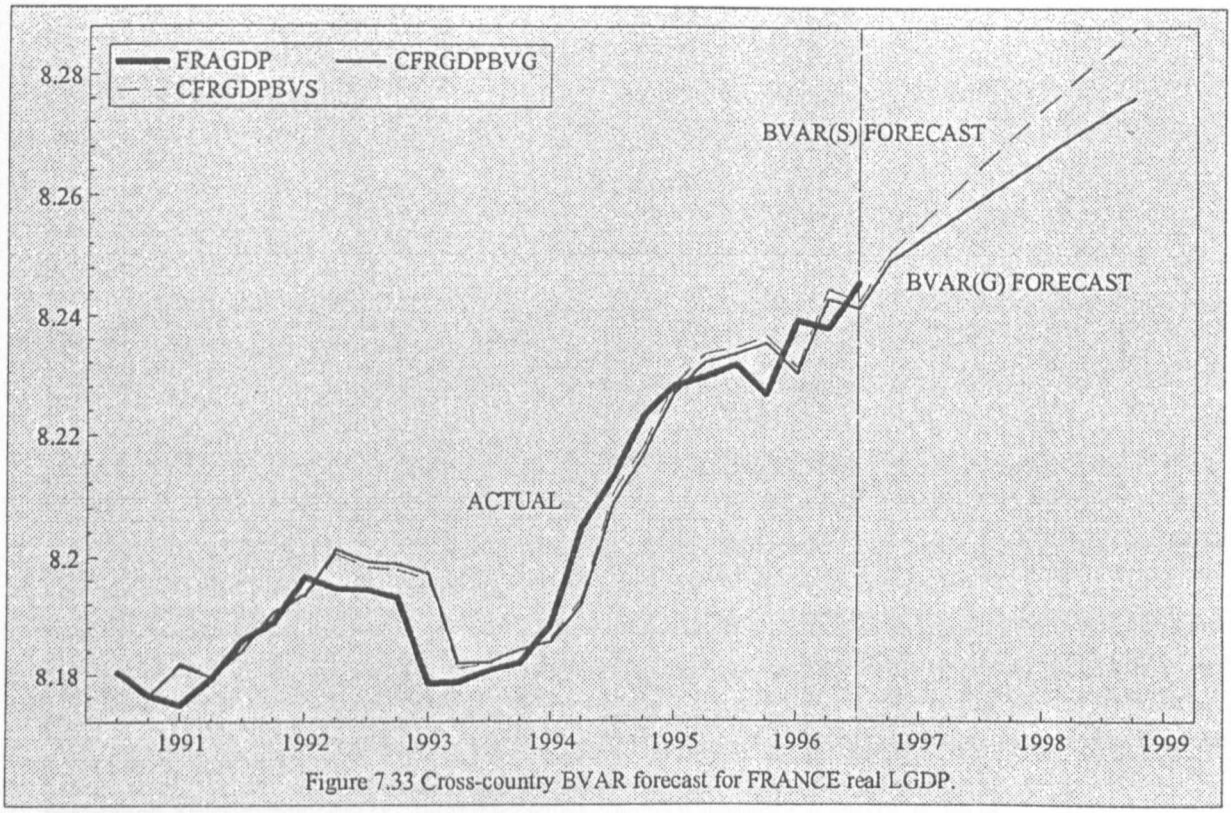

Figure 7.33 Cross-country BVAR forecast for FRANCE real LGDP.

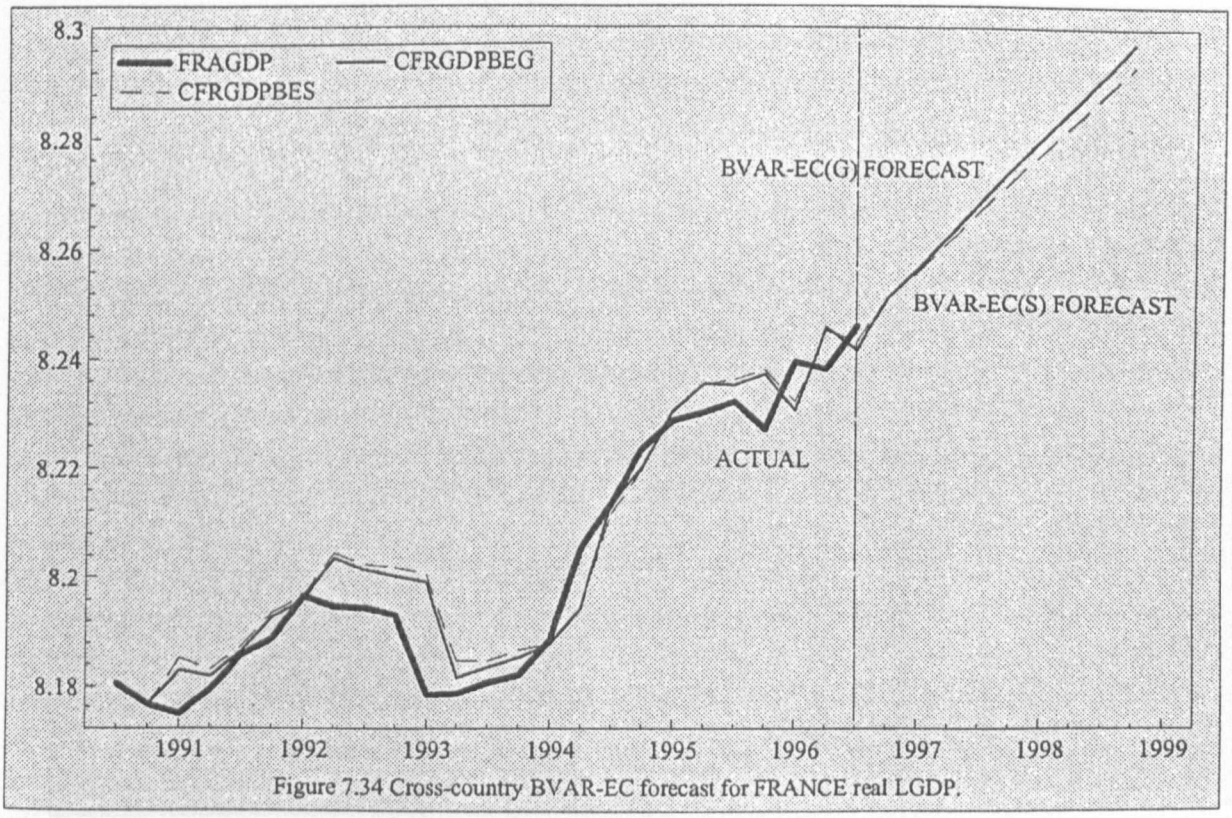

Figure 7.34 Cross-country BVAR-EC forecast for FRANCE real LGDP.

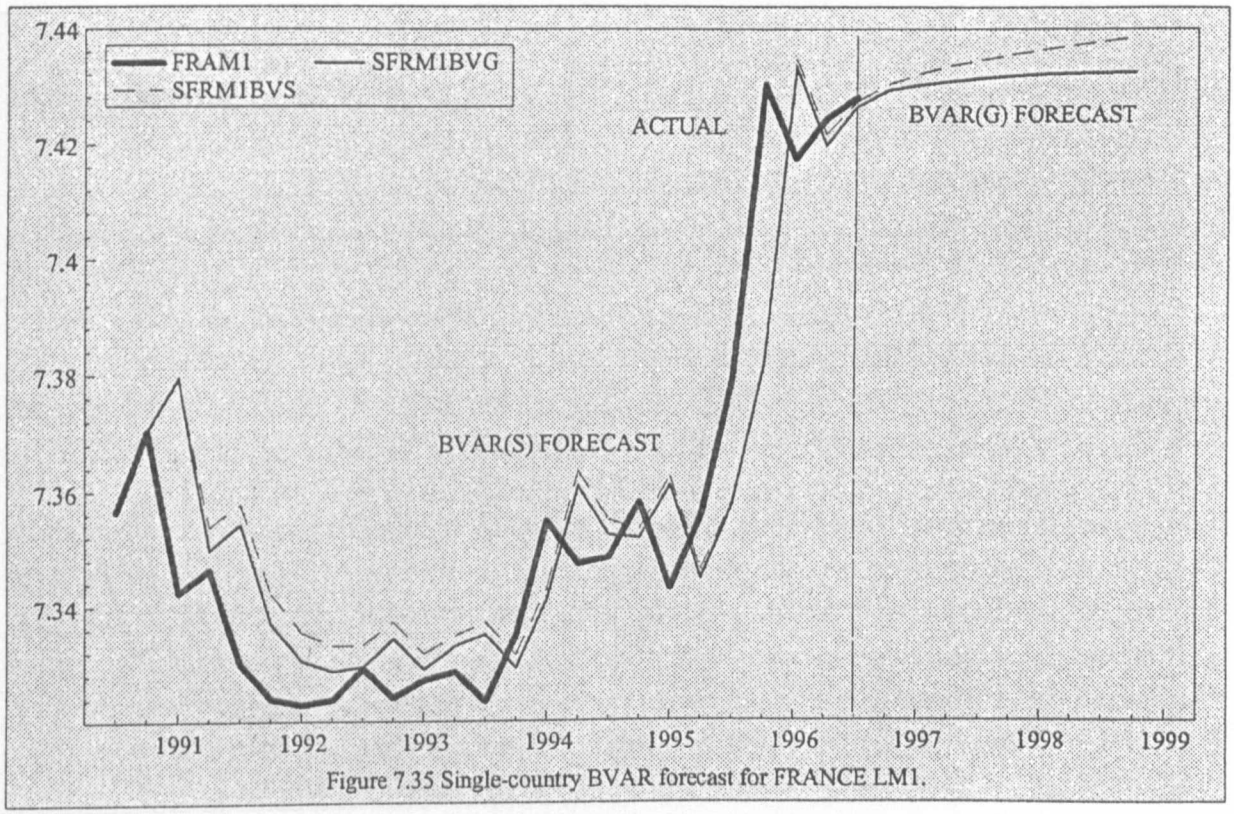

Figure 7.35 Single-country BVAR forecast for FRANCE LM1.

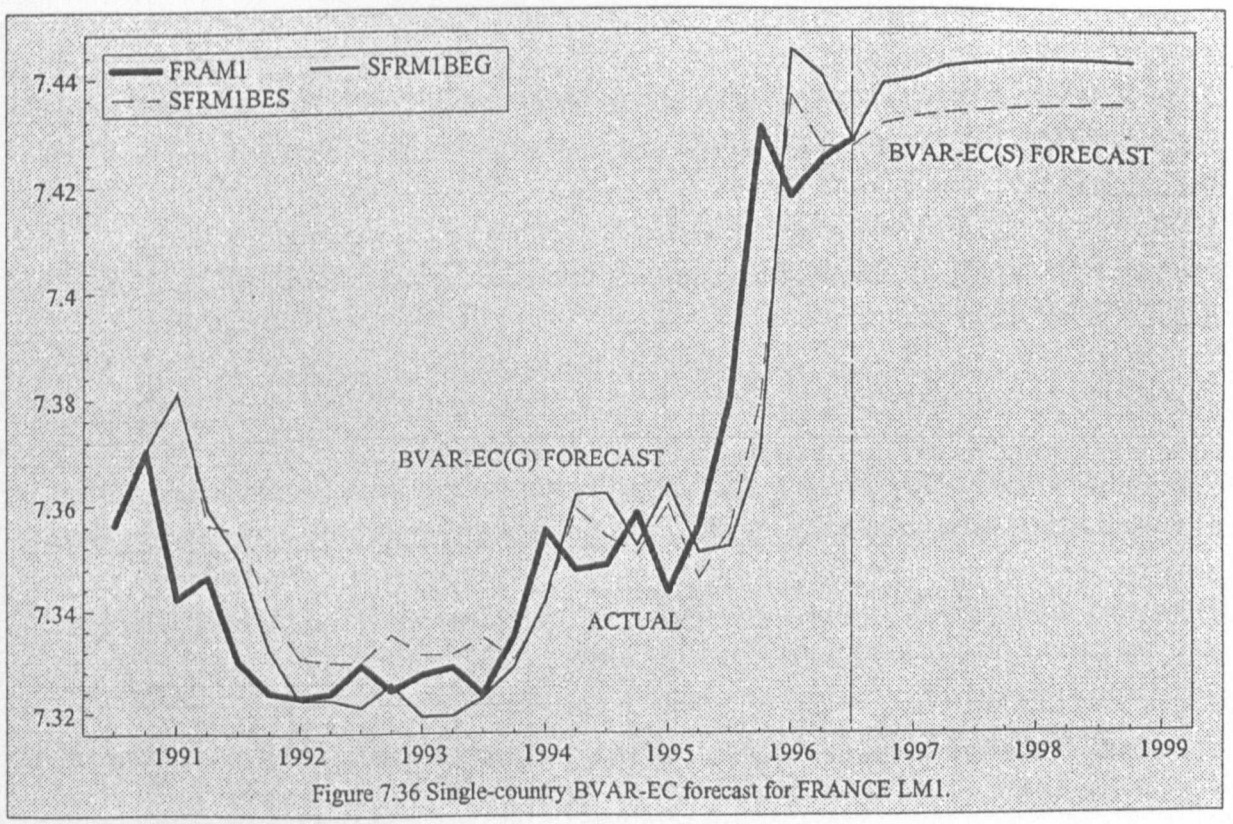

Figure 7.36 Single-country BVAR-EC forecast for FRANCE LM1.

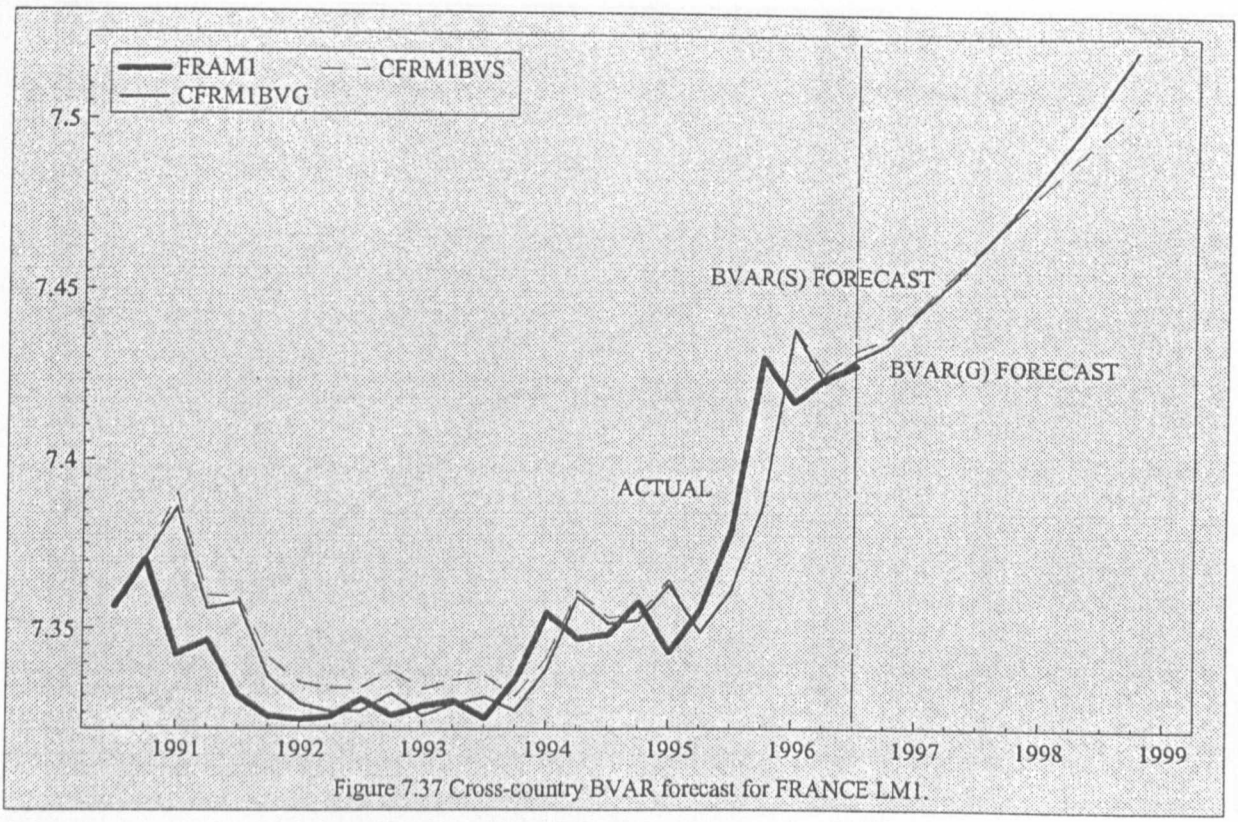

Figure 7.37 Cross-country BVAR forecast for FRANCE LM1.

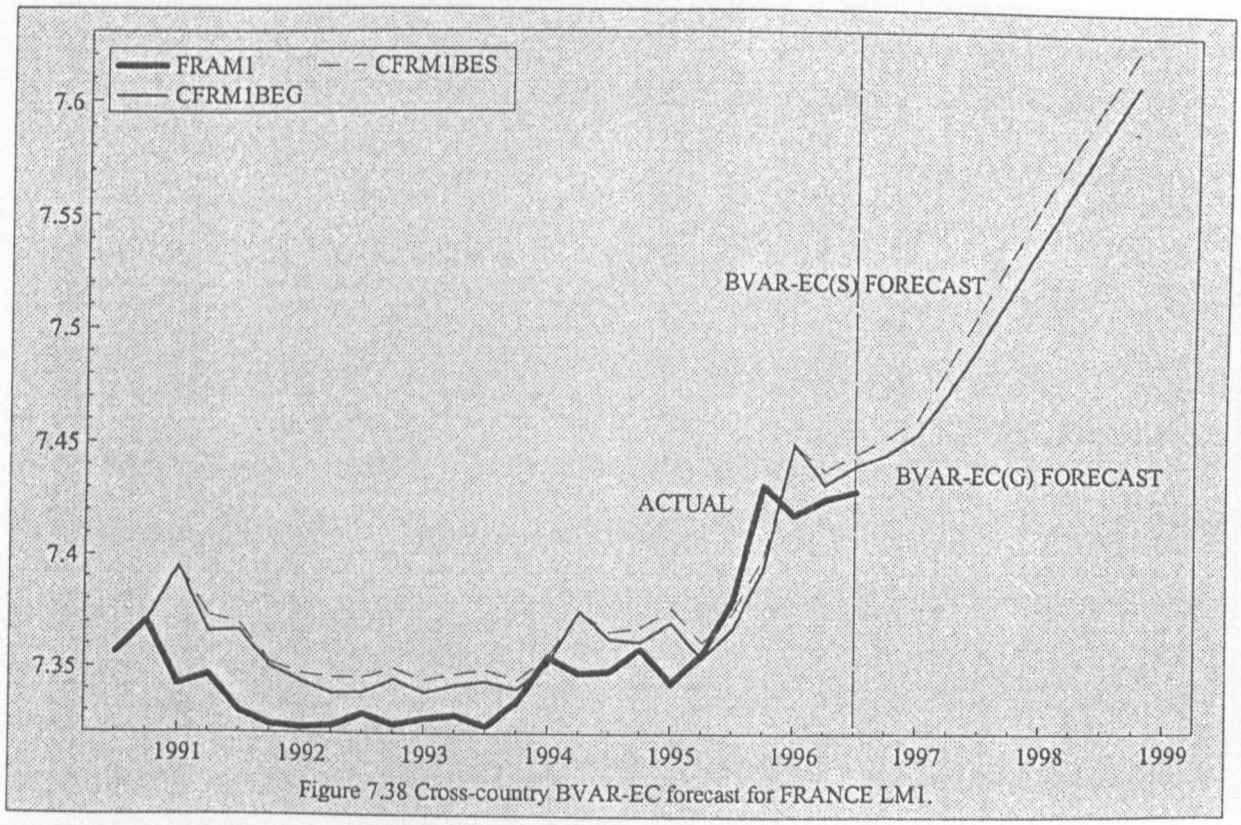

Figure 7.38 Cross-country BVAR-EC forecast for FRANCE LM1.

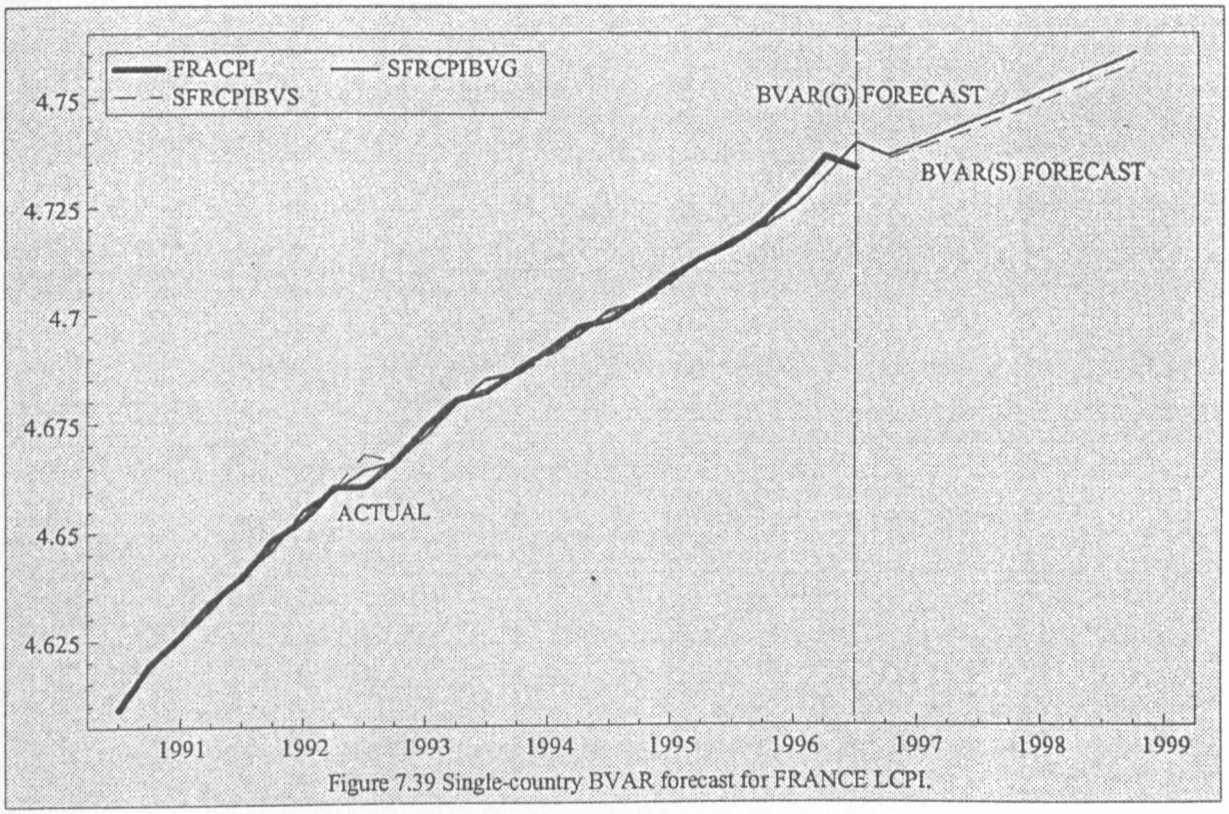

Figure 7.39 Single-country BVAR forecast for FRANCE LCPI.

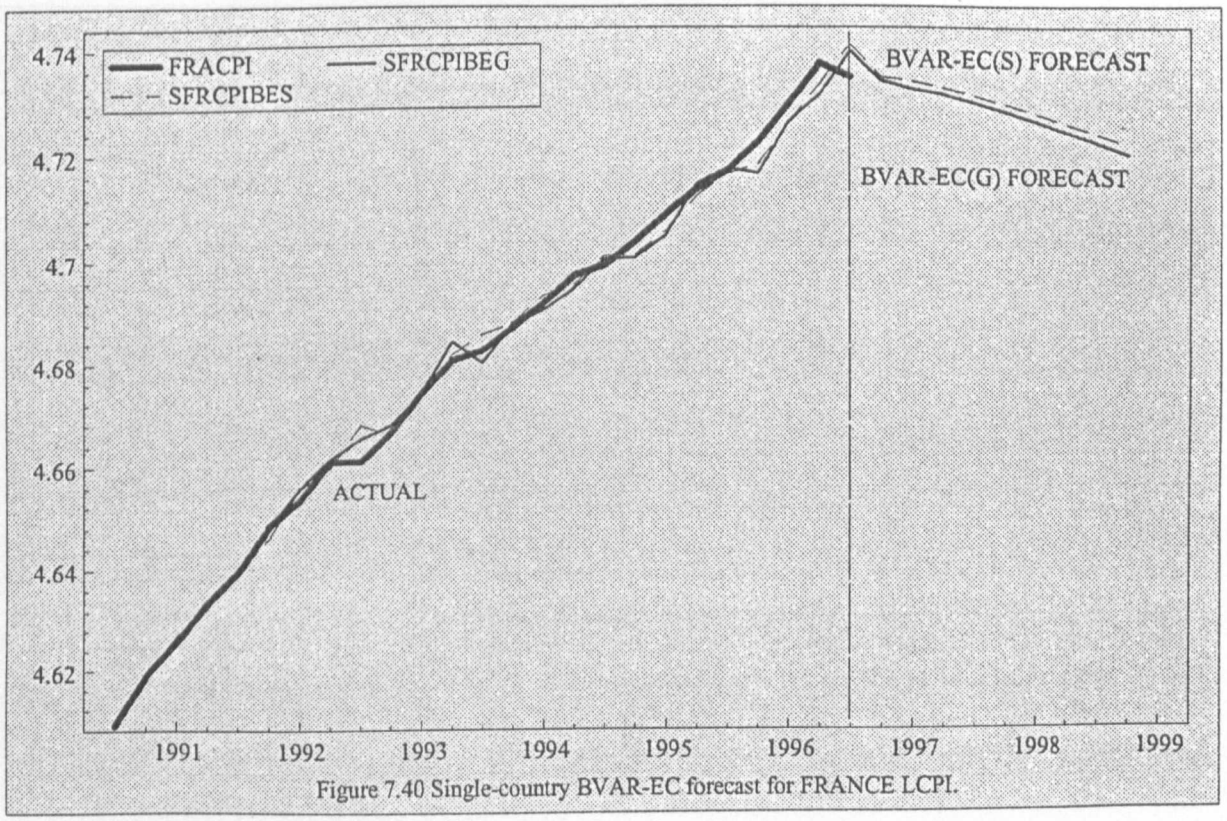

Figure 7.40 Single-country BVAR-EC forecast for FRANCE LCPI.

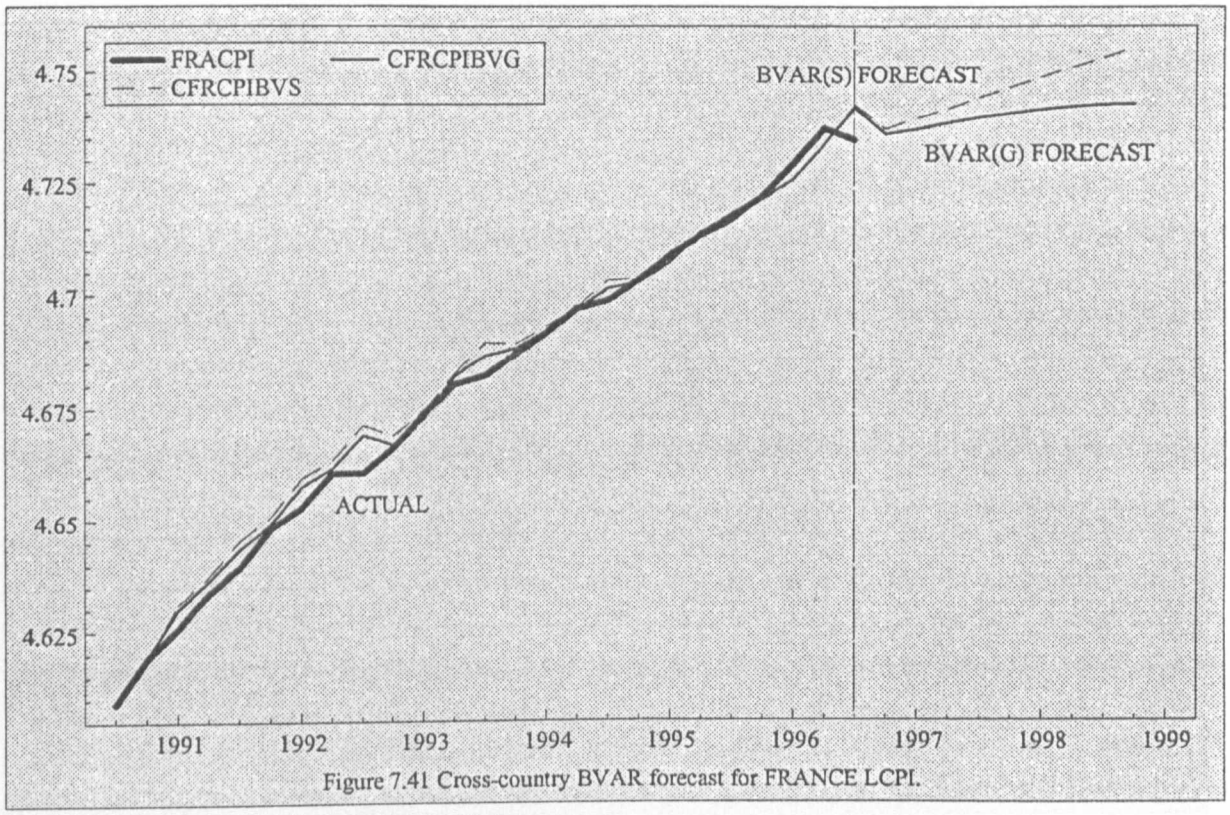

Figure 7.41 Cross-country BVAR forecast for FRANCE LCPI.

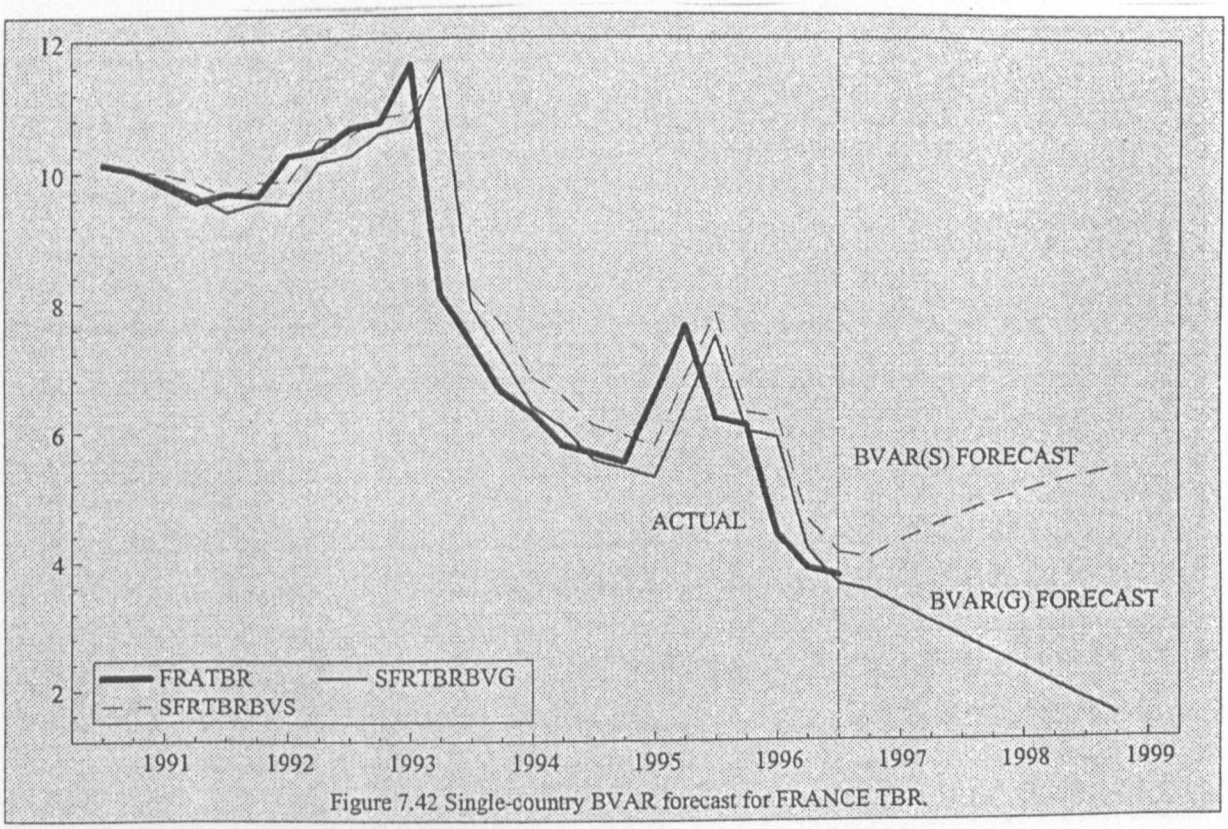

Figure 7.42 Single-country BVAR forecast for FRANCE TBR.

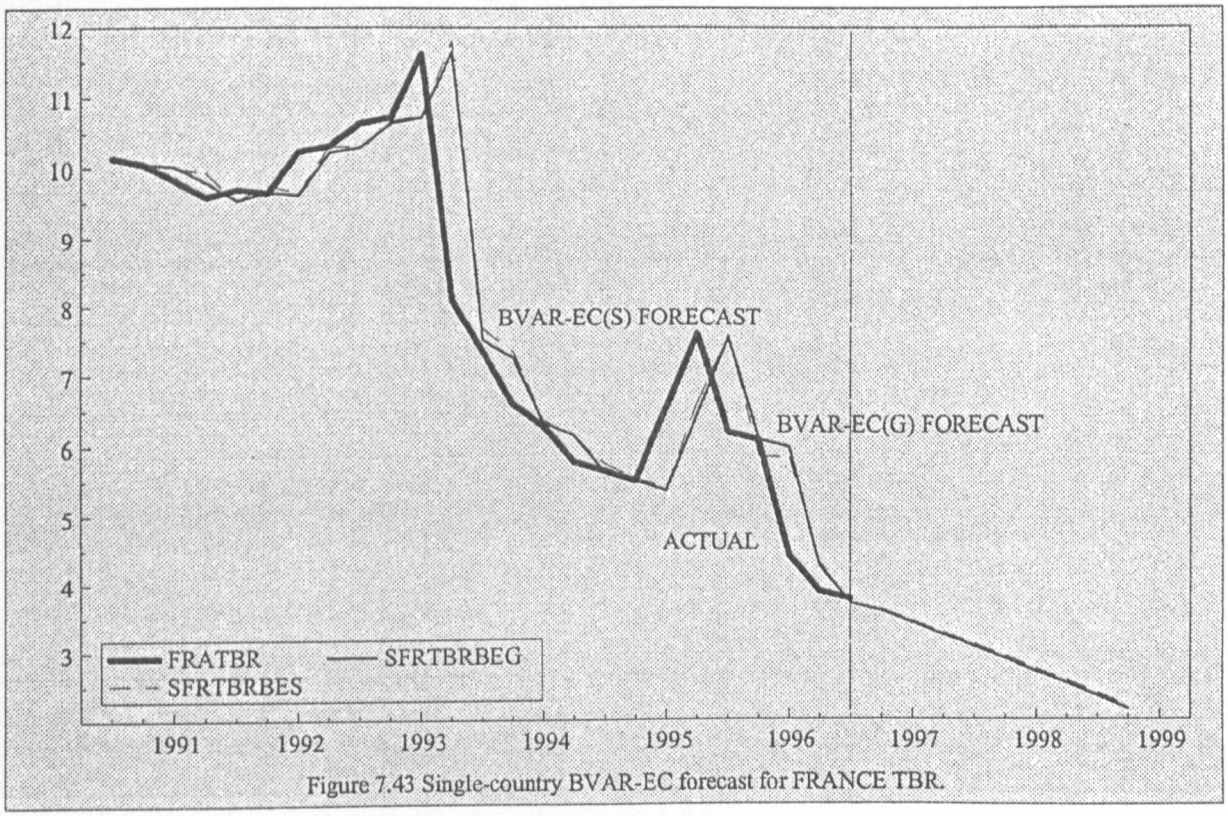

Figure 7.43 Single-country BVAR-EC forecast for FRANCE TBR.

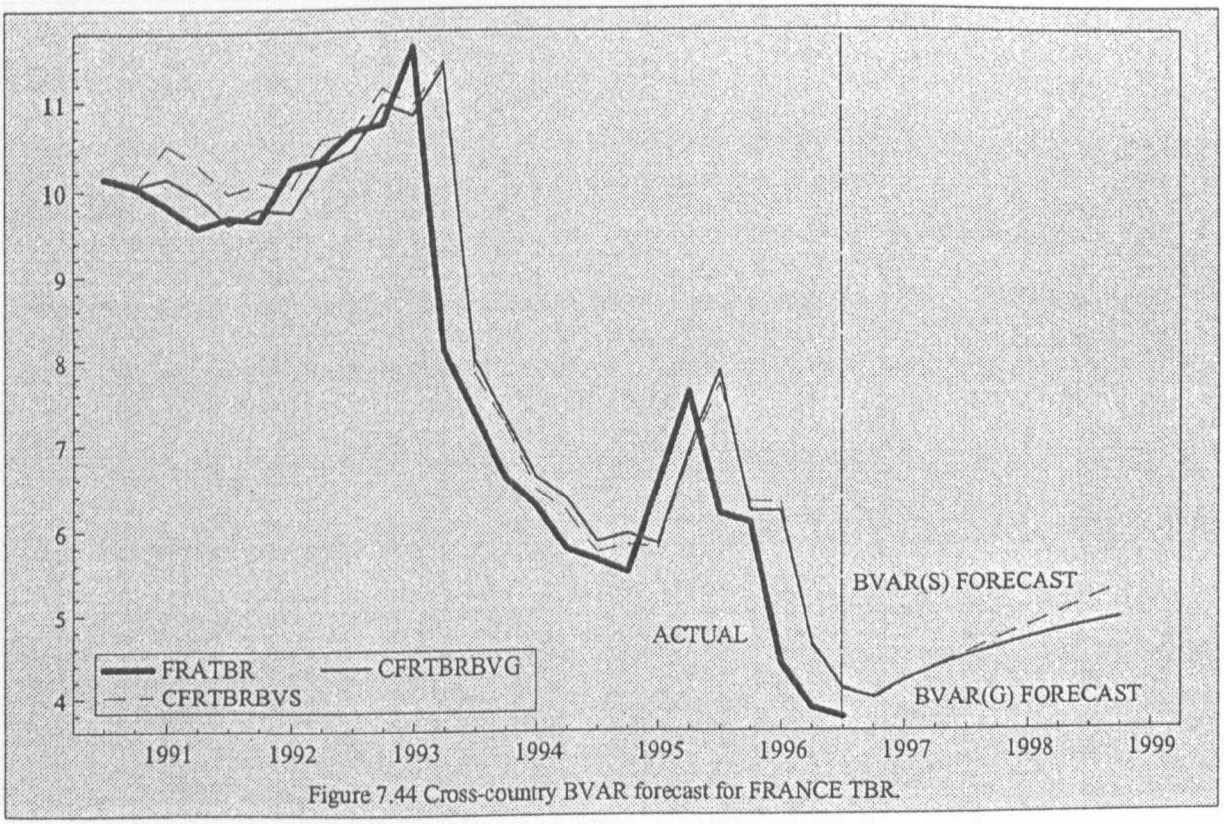

Figure 7.44 Cross-country BVAR forecast for FRANCE TBR.

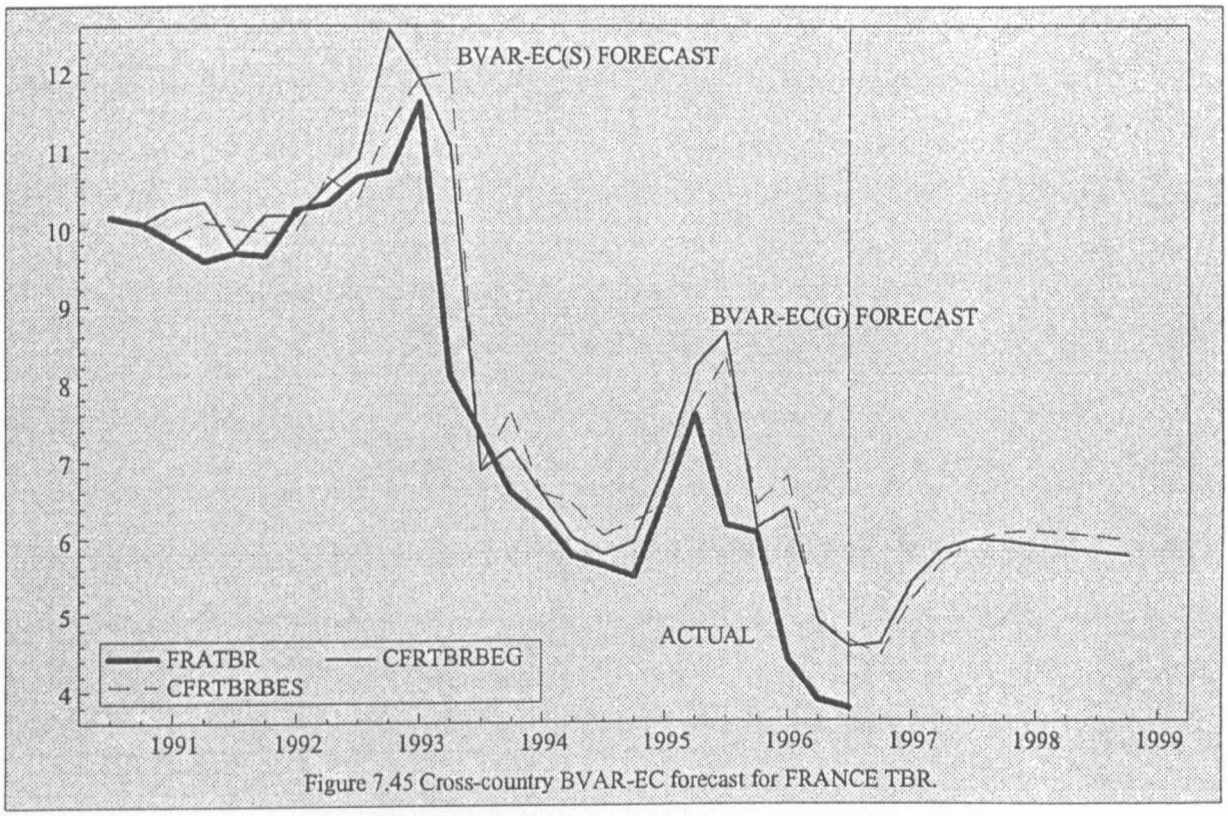

Figure 7.45 Cross-country BVAR-EC forecast for FRANCE TBR.

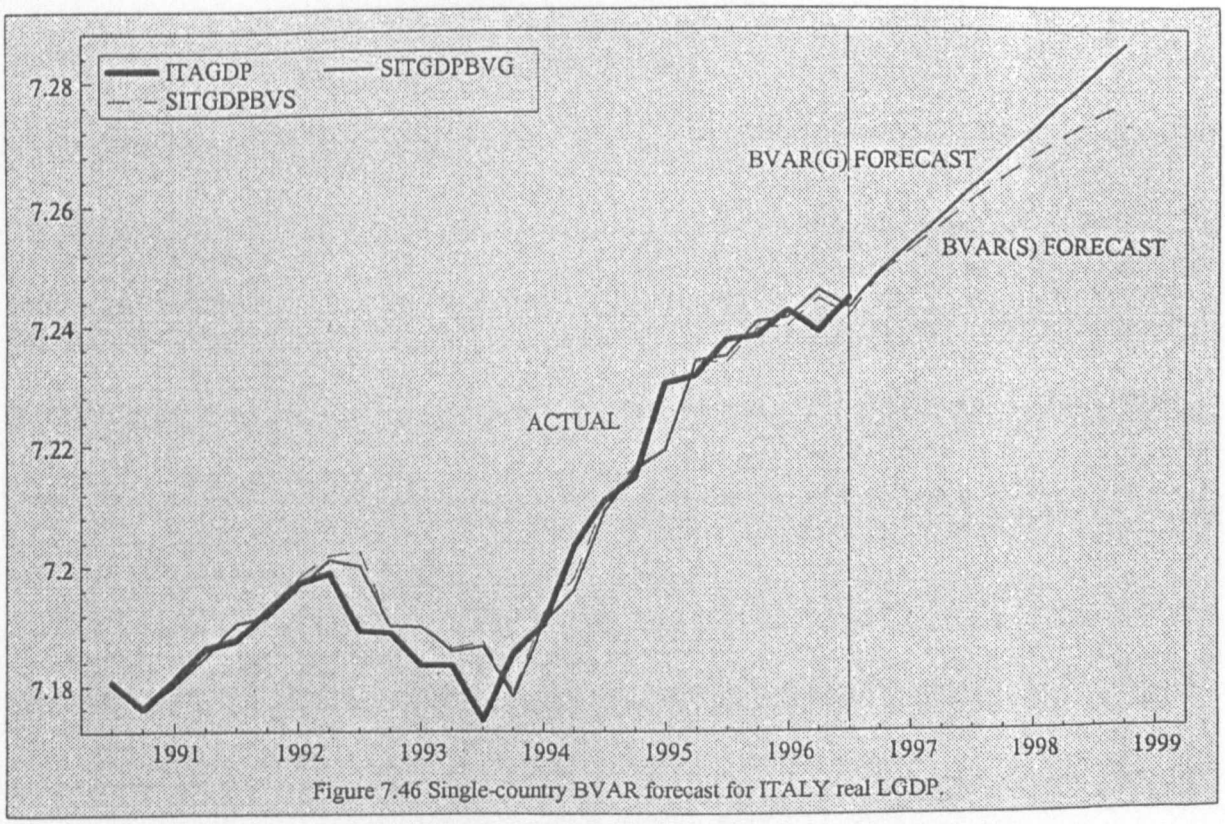

Figure 7.46 Single-country BVAR forecast for ITALY real LGDP.

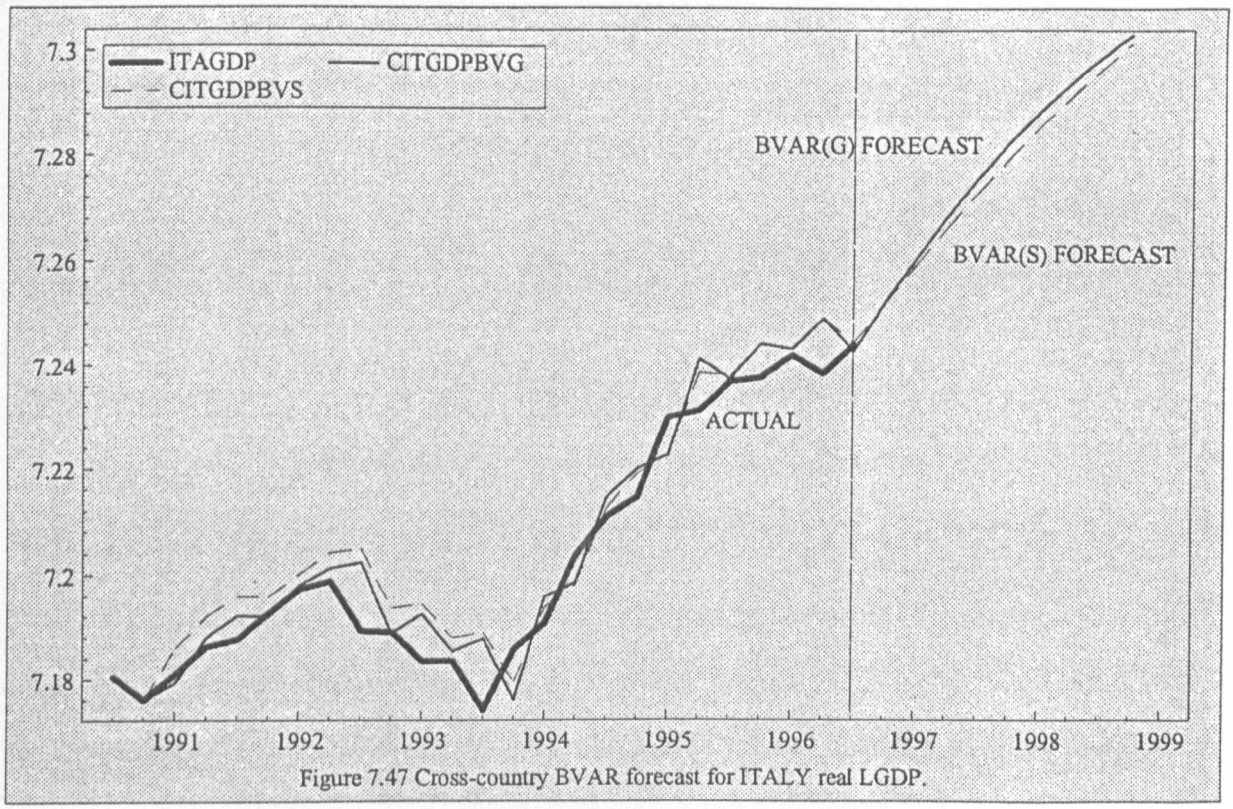

Figure 7.47 Cross-country BVAR forecast for ITALY real LGDP.

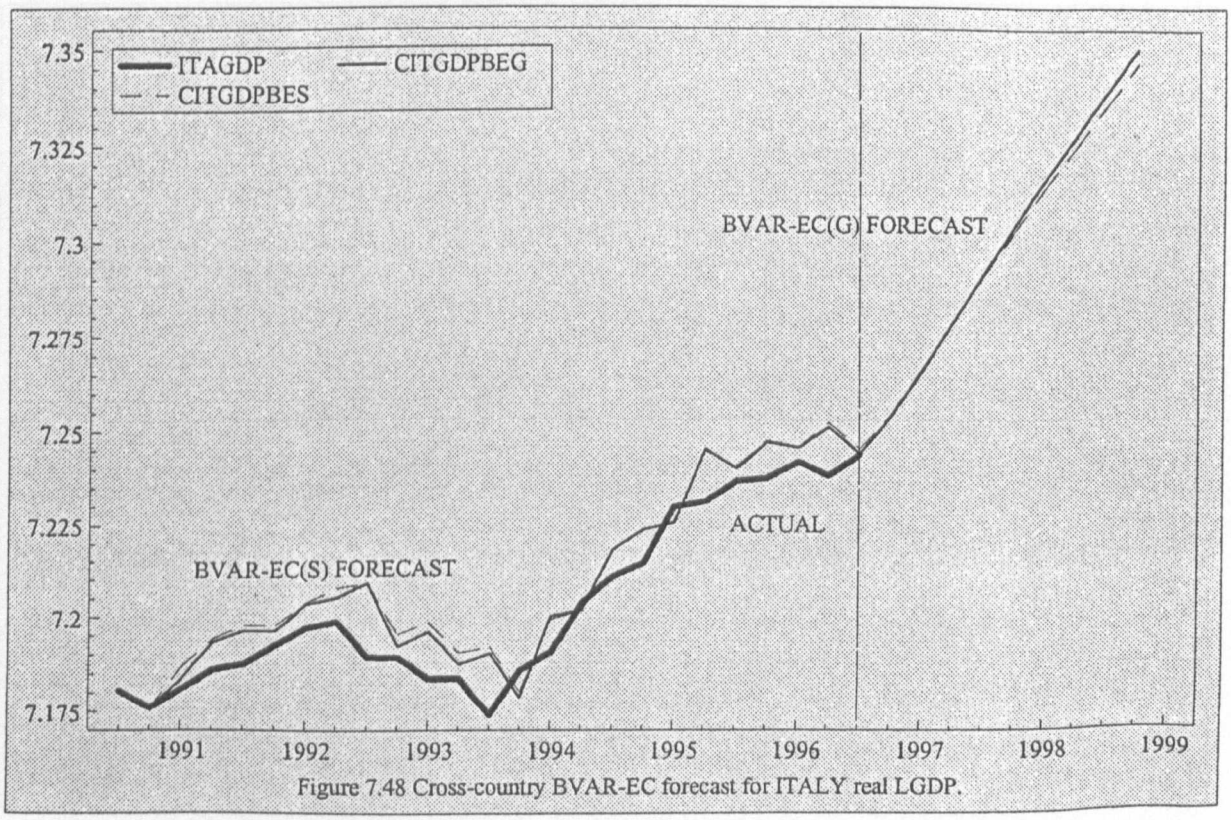

Figure 7.48 Cross-country BVAR-EC forecast for ITALY real LGDP.

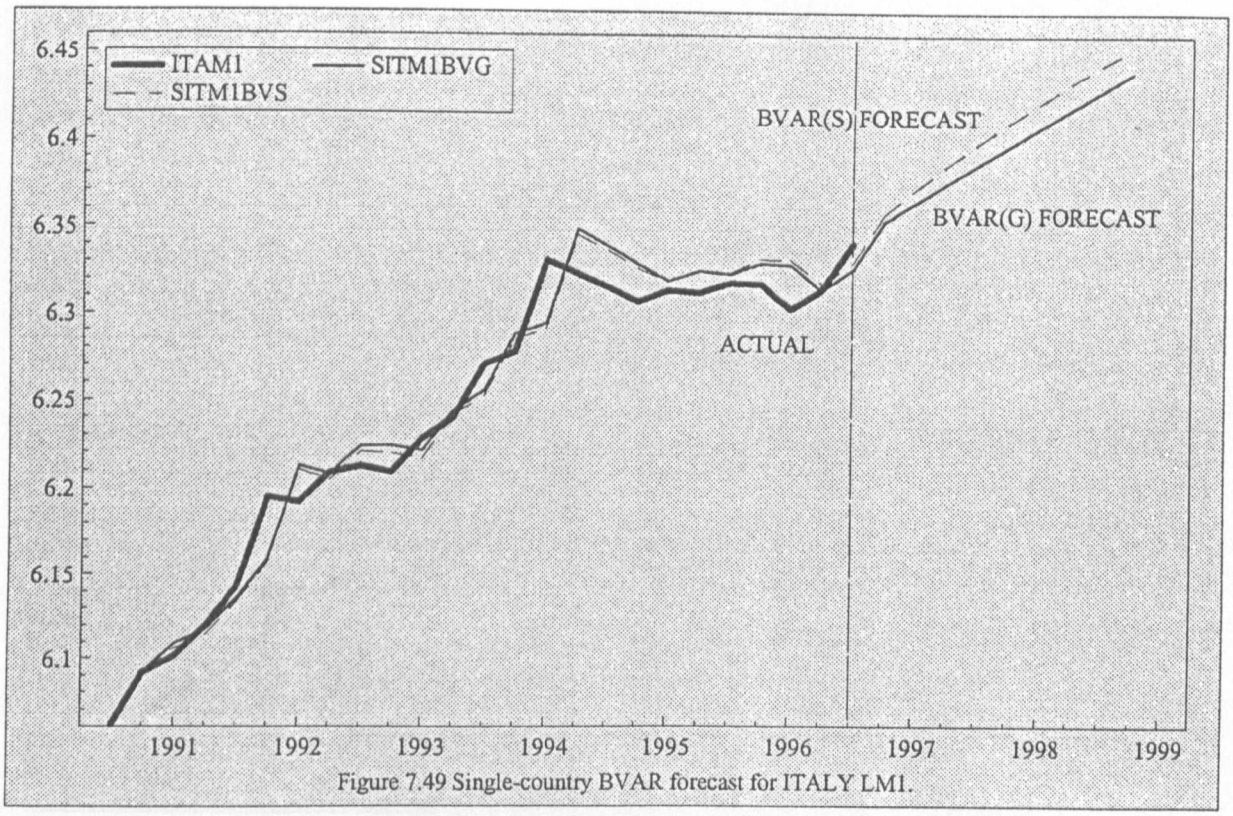

Figure 7.49 Single-country BVAR forecast for ITALY LM1.

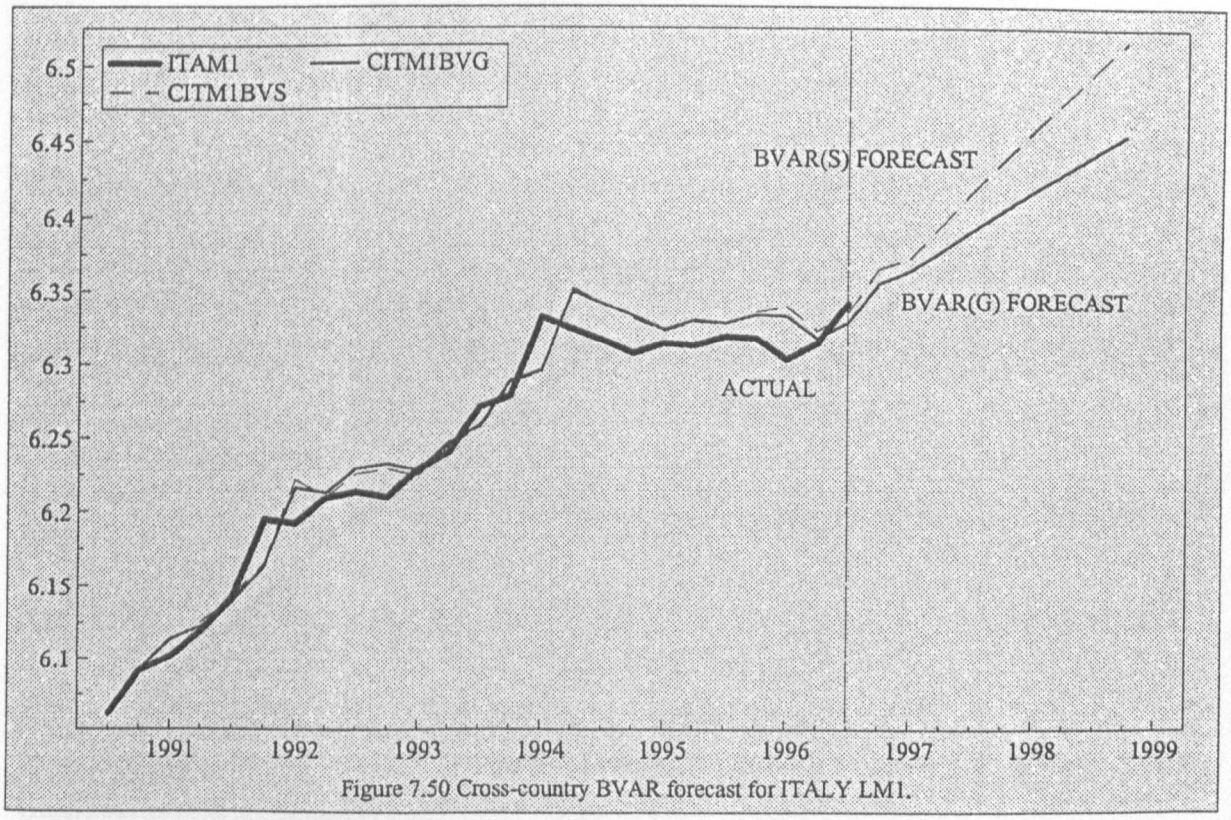

Figure 7.50 Cross-country BVAR forecast for ITALY LM1.

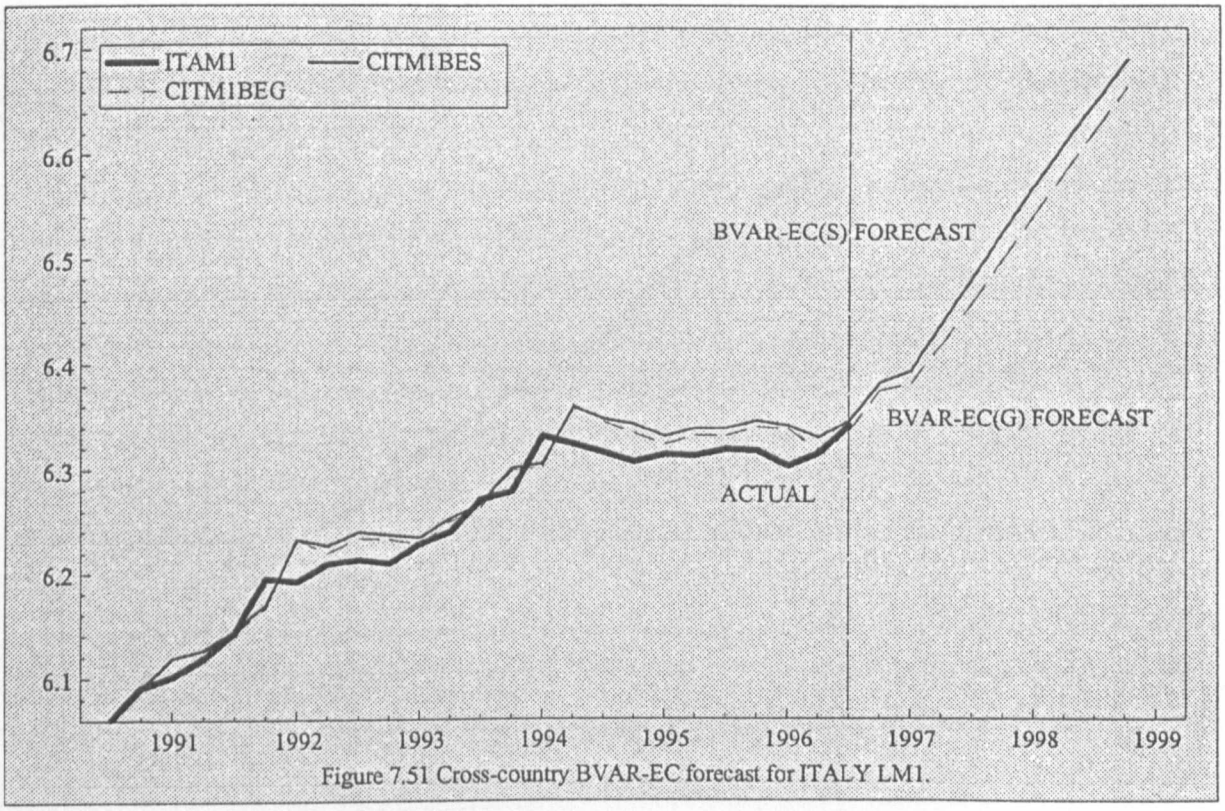

Figure 7.51 Cross-country BVAR-EC forecast for ITALY LM1.

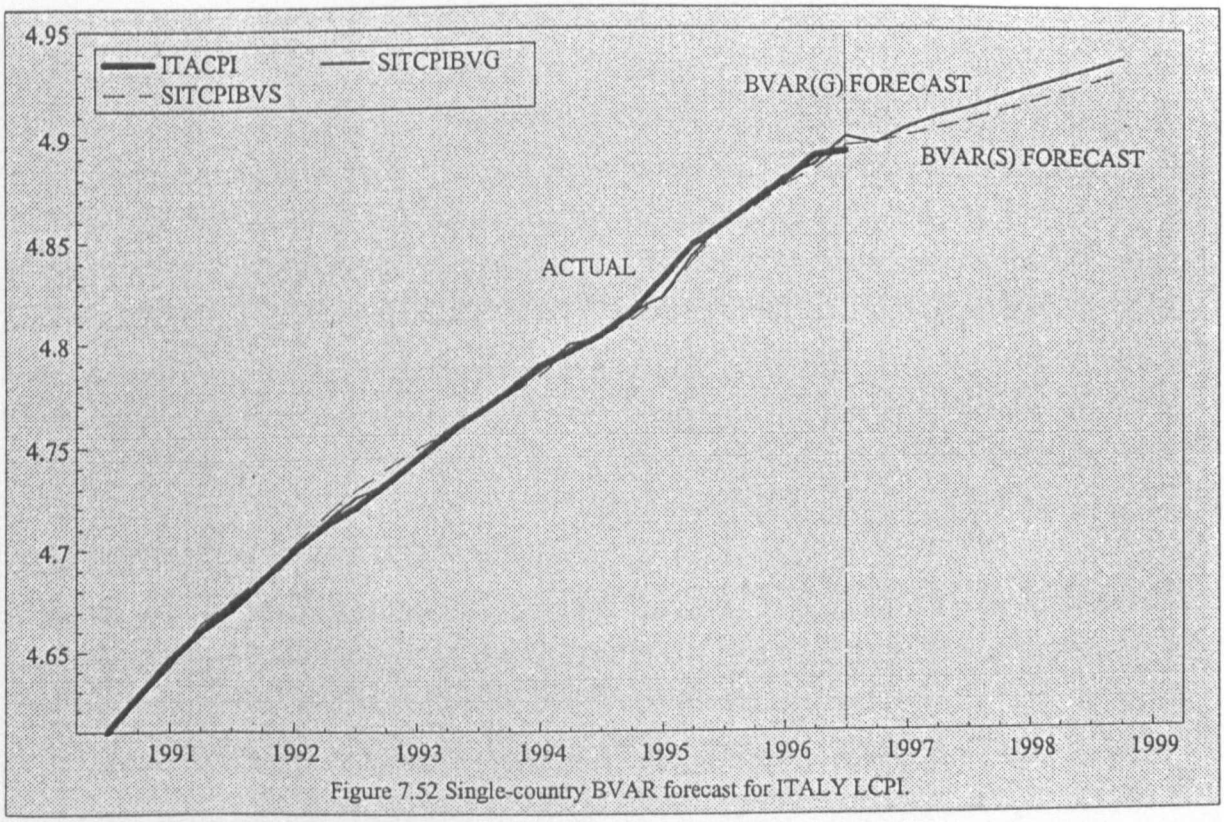

Figure 7.52 Single-country BVAR forecast for ITALY LCPI.

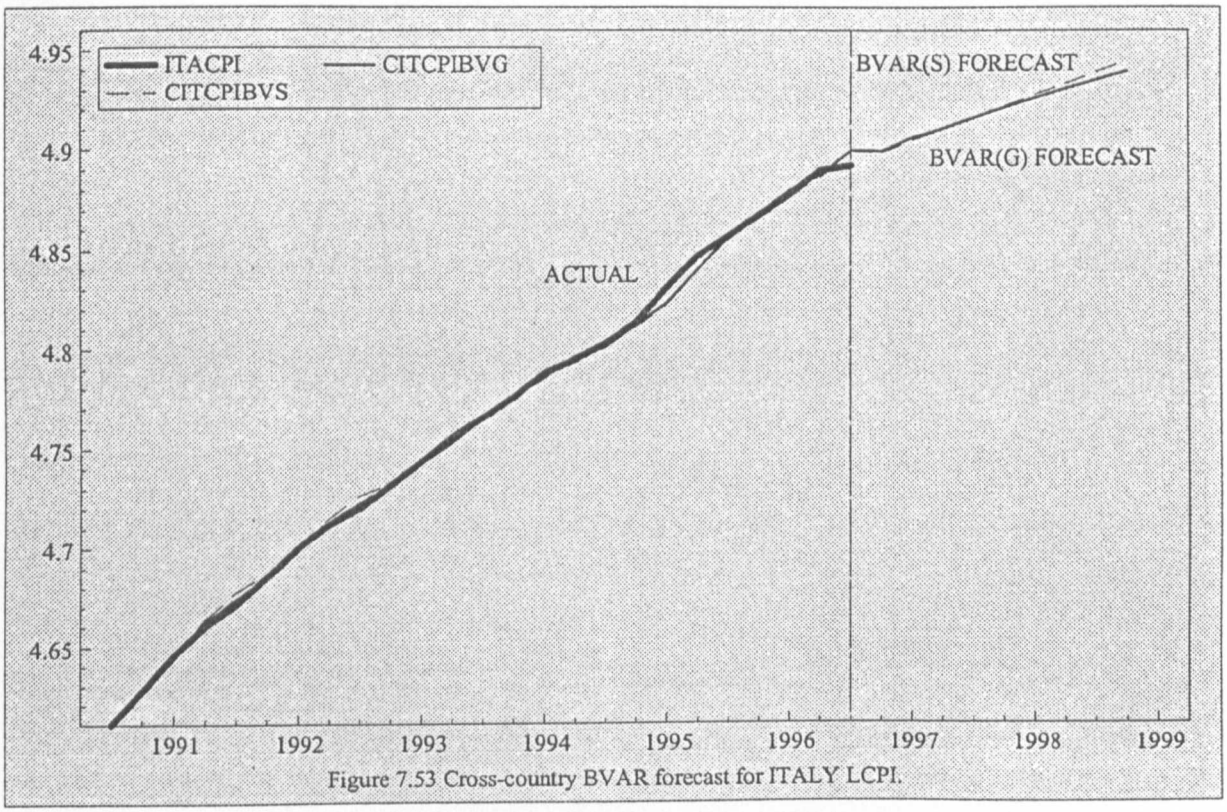

Figure 7.53 Cross-country BVAR forecast for ITALY LCPI.

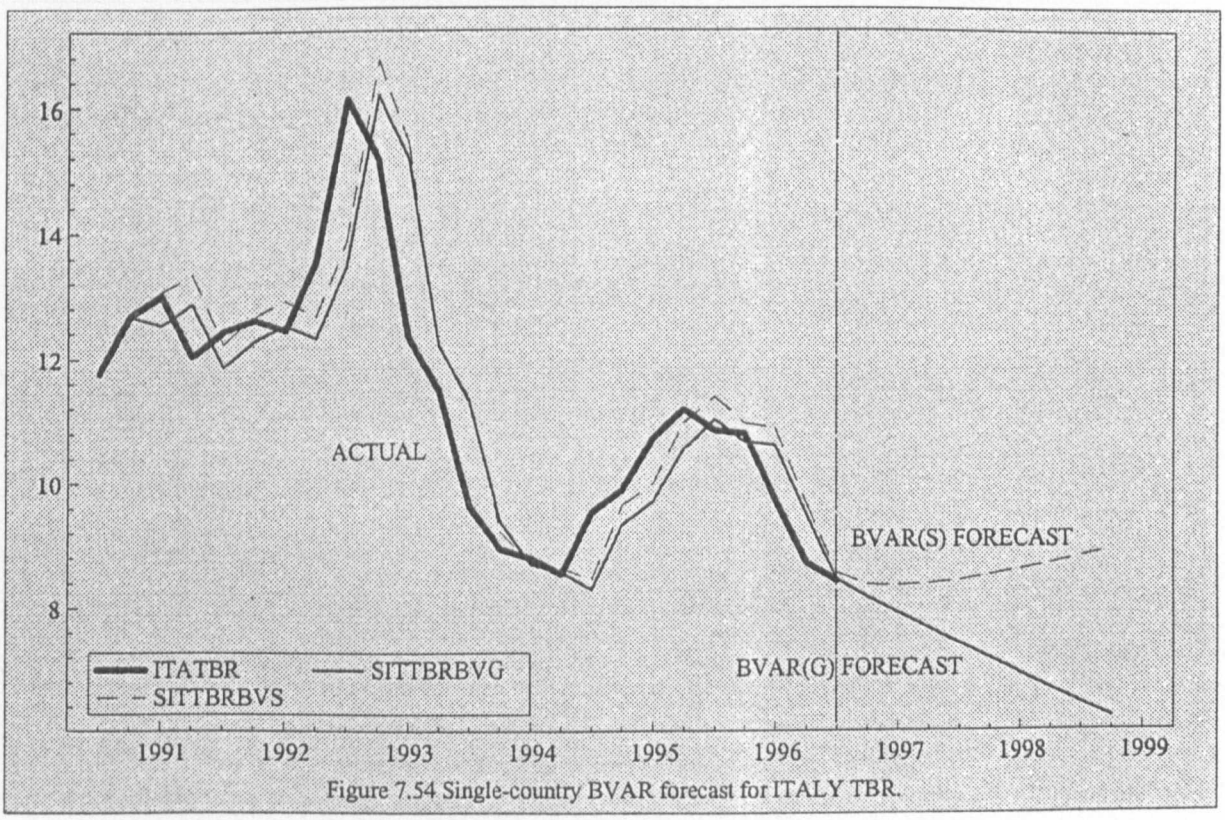

Figure 7.54 Single-country BVAR forecast for ITALY TBR.

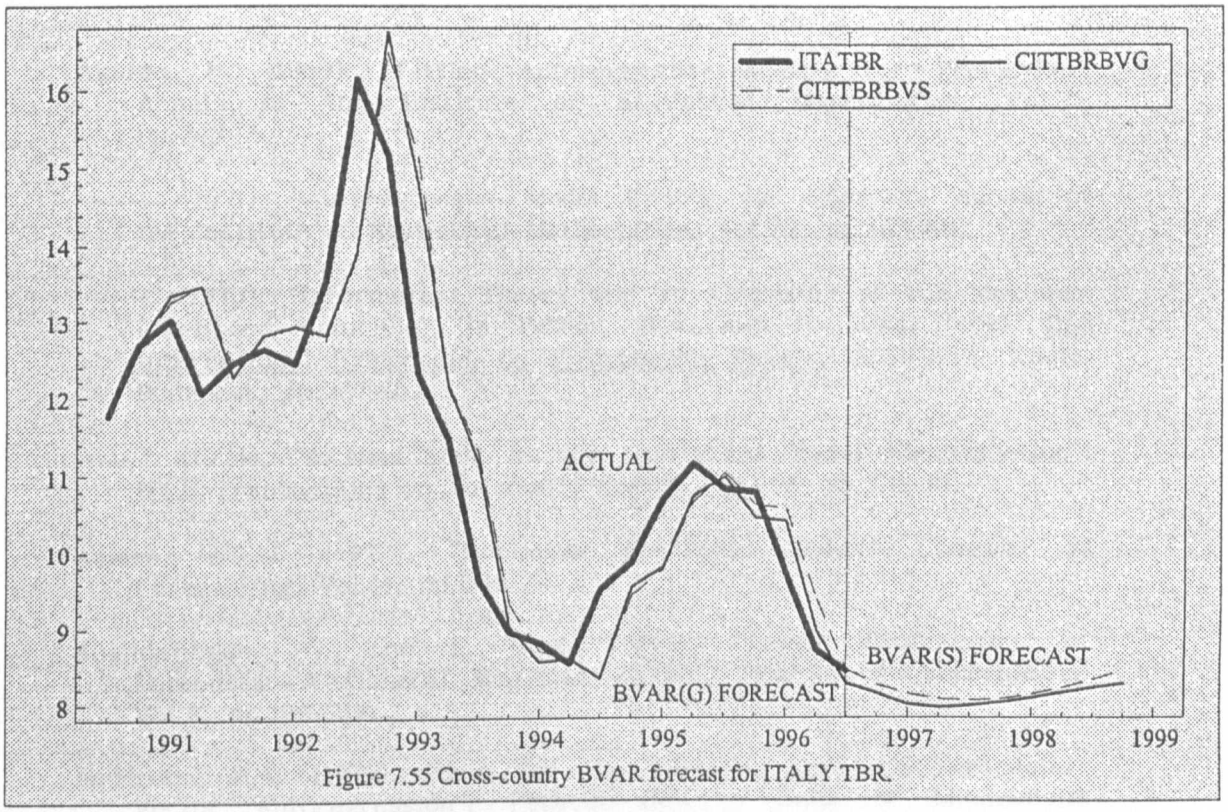

Figure 7.55 Cross-country BVAR forecast for ITALY TBR.

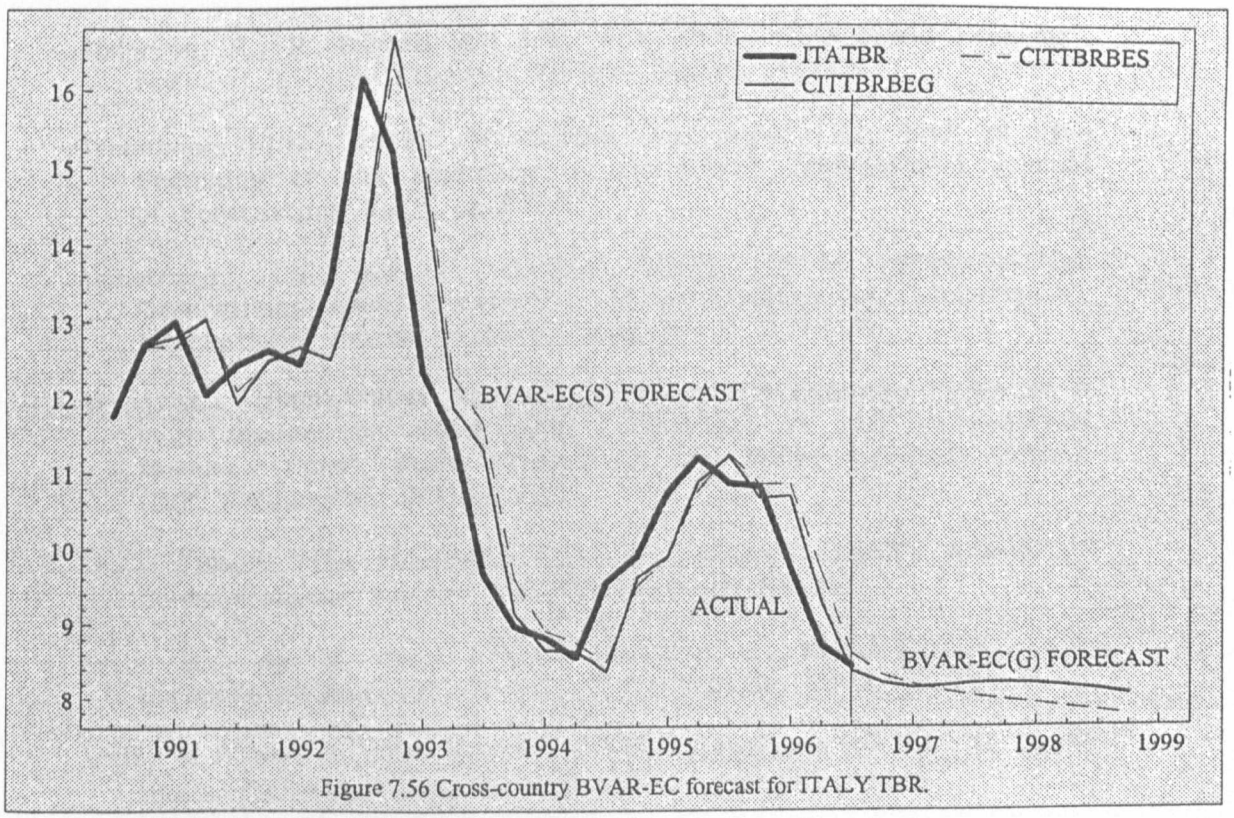

Figure 7.56 Cross-country BVAR-EC forecast for ITALY TBR.

REFERENCES

Akaike, H. (1969), "Fitting Autoregressive Models for Prediction", Annals of the Institute of Statistical Mathematics, Vol.21, pp.243-7.

—— (1971), "Autoregressive Model Fitting for Control", Annals of the Institute of Statistical Mathematics, Vol.23, pp.163-80.

—— (1973), "Information Theory and an Extension of the Maximum Likelihood Principle", in Petrov, B.N. and F. Csáki (eds), 2nd International Symposium on Information Theory, Académiai Kiadó, Budapest, pp.267-81.

—— (1974), "A New Look at the Statistical Model Identification", IEEE Transactions on Automatic Control, AC-19, pp.716-23.

Ameen, J.R.M. (1992), "Nonlinear Predictor Models", Journal of Forecasting, Vol.11, pp.309-24.

Anandalingam, G. and L. Chen (1989), "Linear Combination of Forecasts: A General Bayesian Model", Journal of Forecasting, Vol.8, pp.199-214.

Anderson, P.A. (1979), "Help for the Regional Economic Forecaster: Vector Autoregression", Federal Reserve Bank of Minneapolis Quarterly Review, Vol.3, No.3, pp.2-7.

Anderson, R., R. Barrell and J.W. Veld (1992), "The World Economy", National Institute Economic Review, August, pp.28-47.

Armstrong, J.S. (1989), "Combining Forecasts: The End of the Beginning or the Beginning of the End?", International Journal of Forecasting, Vol.5, pp.585-8.

Armstrong, J.S. and F. Collopy (1992), "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons", International Journal of Forecasting, Vol.8, pp.69-80.

Artis, M.J. (1988), "How Accurate Is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund", Staff Studies for the World Economic Outlook (IMF, Washington, DC), pp.1-49.

—— (1992), "The Maastricht Road to Monetary Union", Journal of Common Market Studies, Vol.XXX, No.3, pp.299-310.

Artis, M.J. and S. Holly (1992), "Modelling the World Economy: The European Perspective", Journal of Forecasting, Vol.11, pp.333-40.

Artis, M.J. and N. Lee (ed) (1997), The Economics of The European Union, Policy and Analysis (2nd edn), Oxford UP, Oxford.

Artis, M.J. and W.D. Zhang (1990a), "BVAR Forecasts for the G-7", International Journal of Forecasting, Vol.6, pp.349-62.

—— (1990b), "BVAR Forecasts of the World Economy", CEPR Discussion Paper Series, No.380, pp.1-19.

Banerjee, A. (1993), Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data, Oxford UP, Oxford.

Bannock, G., R.E. Baxter and E. Davis (1987), The Penguin Dictionary of Economics (4th edn), Penguin Books Ltd, London.

Barbosa, E. and J. Harrison (1992), "Variance Estimation for Multivariate Dynamic Linear Models", Journal of Forecasting, Vol.11, pp.621-8.

Barker, R., M. Franklin and J. Walker (1993), "Conflicting Signs Cloud the Picture — House Price and Jobless Figures Point Towards a Slump But Other Indicators Hint at Economic Recovery", the Independent on Sunday, 31 January.

Barrell, R.J. and N.C. Pain (1992), "The Comparison and Evaluation of Forecasts of the World Economy", mimeo, National Institute of Economic and Social Research, pp.1-12.

Begg, D.K.H. (1982), The Rational Expectations Revolution in Macroeconomics, Camelot Press Ltd, Southampton.

Begg, D., S. Fischer and R. Dornbusch (1991), Economics (3rd edn), McGraw-Hill Book Company, London.

Bewley, R. (1986), Allocation Models: Specification, Estimation, and Applications, Ballinger Publishing Company, Cambridge.

Bhargava, A. (1986), "On The Theory of Testing for Unit Roots in Observed Time Series", Review of Economic Studies, Vol.LIII, pp.369-84.

Bladen-Hovell, R.C. and W.D. Zhang (1991), "A BVAR Model for the UK Economy: A Forecast Comparison with the LBS and NI Models", Discussion Paper, No.8, ICMM, Glasgow Development Agency, pp.1-31.

Blanchard, O.J. (1987), "Comment", Journal of Business & Economic Statistics, Vol.5, No.4, pp.449-51.

—— (1989), "A Traditional Interpretation of Macroeconomic Fluctuations", American Economic Review, Vol.79, No.5, pp.1146-64.

Blanchard, O.J. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", American Economic Review, Vol.79, No.4, pp.655-73.

Blinder, A.S. and S. Fischer (1981), "Inventories, Rational Expectations, and the Business Cycle", Journal of Monetary Economics, Vol.8, pp.277-304.

Bodkin, R.G., L.R. Klein and K. Marwah (eds) (1991), A History of Macroeconometric Model-Building, Billing & Sons Ltd., Worcester.

Boswijk, H.P. (1995), "Efficient Inference on Cointegration Parameters in Structural Error Correction Models", Journal of Econometrics, Vol.69, pp.133-58.

Box, G.E.P. and G.M. Jenkins (1970), Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco.

—— (1976), Time Series Analysis: Forecasting and Control (Revised edn.), Holden-Day, San Francisco.

Brandner, P. and R.M. Kunst (1990), "Forecasting Vector Autoregressions — the Influence of Cointegration: A Monte Carlo Study", Research Memorandum No.265, Institute for Advanced Studies, Vienna.

Branson, W.H. (1989), Macroeconomic Theory and Policy (3rd edn), Harper & Row, New York.

Braun, P.A. and S. Mittnik (1993), "Misspecifications in Vector Autoregressions and Their Effects on Impulse Responses and Variance Decompositions", Journal of Econometrics, Vol.59, pp.319-41.

Brockwell, P.J. and R.A. Davis (1991), Time Series: Theory and Methods (2nd edn), Springer Verlag, New York.

Campos, J., N.R. Ericsson and D.F. Hendry (1996), "Cointegration Tests in the Presence of Structural Breaks", Journal of Econometrics, Vol.70, pp.187-220.

Capie, F.H. and T.C. Mills (1991), "Money and Business Cycles in the U.S. and U.K., 1870 to 1913", the Manchester School, pp.38-56.

Charemza, W. and D.F. Deadman (1992), New Directions in Econometric Practice: General to Specific Modelling, Cointegration, and Vector Autoregression, E. Elgar, Aldershot.

Chiang, A.C. (1984), Fundamental Methods of Mathematical Economics (3rd edn), McGraw-Hill, London.

Chib, S. (1993), "Bayes Regression with Autoregressive Errors - A Gibbs Sampling Approach", Journal of Econometrics, Vol.58, pp.275-94.

Chong, Y.Y. and D.F. Hendry (1986), "Econometric Evaluation of Linear Macro-Economic Models", Review of Economic Studies, Vol.LIII, pp.671-90.

Chow, C.C. and A.L. Lin (1971), "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series", Review of Economics and Statistics, Vol.53, pp.372-5.

Christiano, L. and M. Eichenbaum (1989), "Unit Roots in Real GNP: Do

We Know and Do We Care?", Discussion Paper, Northwestern University.

Clemen, R.T. (1989), "Combining Forecasts: A Review and Annotated Bibliography", International Journal of Forecasting, Vol.5, pp.559-83.

Clements, M.P. (1996), "Evaluating the Rationality of Fixed-Event Forecasts", Warwick Economic Research Papers, No.457, pp.1-16.

Clements, M.P. and D.F. Hendry (1992a), "Towards a Theory of Economic forecasting", Discussion Paper, Institute of Economics and Statistics and Nuffield College, Oxford, pp.1-38.

—— (1992b), "On the Limitations of Comparing Mean Square Forecast Errors", Applied Economics Discussion Paper Series, No.138, pp.1-19.

—— (1995), "Forecasting in Cointegrated Systems", Journal of Applied Econometrics, Vol.10, pp.127-46.

—— (1996), Multi-Step Estimation for Forecasting", Warwick Economic Research Papers, No.447, pp.1-32.

Clements, M.P. and H.M. Krolzig (1997), "A Comparison of the Forecast Performance of Markov-Switching and Threshold Autoregressive Models of US GNP", Warwick Economic Research Papers, No.489, pp.1-29.

Clements, M.P. and R. Madlener (1997), "Seasonality, Cointegration, and the Forecasting of Energy Demands", Warwick Economic Research Papers, No.484. pp.1-34.

Clements, M.P. and G.E. Mizon (1991), "Empirical Analysis of Macroeconomic Time Series — VAR and Structural Models", European Economic Review, Vol.35, pp.887-932.

Clements, M.P. and J. Smith (1996a), "A Mote Carlo Study of the Forecasting Performance of Empirical SETAR Models", Warwick Economic Research Papers, No.464, pp.1-75.

—— (1996b), "Performance of Alternative Forecasting Methods for SETAR Models", Warwick Economic Research Papers, No.467, pp.1-32.

—— (1997), "Forecasting Seasonal UK Consumption Components", Warwick Economic Research Papers, No.479, pp.1-20.

Coen, P.J., E.D. Gomme and M.G. Kendall (1969), "Lagged Relationships in Economic Forecasting", Scientific Control Systems Ltd, London, pp.133-63.

Cooley, T.F. and S.F. LeRoy (1985), "Atheoretical Macroeconometrics: a Critique", Journal of Monetary Economics, Vol.16, pp.283-308.

Crafts, N.F.R., S.J. Leybourne and T.C. Mills (1989a), "Trend and Cycles in British Industrial Production, 1700-1913", Journal of Royal Statistical Society, Vol.152, Part 1, pp.43-60.

—— (1989b), "The Climactic in Late Victorian Britain and France: A Reappraisal of the Evidence", Journal of Applied Econometrics, Vol.4, pp.103-17.

Cramer, J.S. (1986), Econometric Applications of Maximum Likelihood Methods, Cambridge UP, Cambridge.

Cubitt, R.P. (1993), "On the Possibility of Rational Dilemmas", Economics and Philosophy, Vol.9, pp.1-23.

Davidson, J.E.H., D.F. Hendry, F. Srba and S. Yeo (1978), "Econometric Modelling of the Aggregate Time Series Relationship between Consumers' Expenditure and Income in the United Kingdom", Economic Journal, Vol.88, pp.661-92.

Davidson, R. and J.G. Mackinnon (1993), Estimation and Inference in Econometrics, Oxford UP, Oxford.

Delessy, H. and H. Harasty (1992), "Oil Shocks: How Can OECD Countries Manage Them Best?", Journal of Forecasting, Vol.11, pp.491-506.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", Journal of the American Statistical Association, Vol.74, No.366, pp.427-32.

—— (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", Econometrica, Vol.49, No.4, pp.1057-72.

Dolado, J.J., T. Jenkinson and S. Sosvilla-Rivero (1990), "Cointegration and Unit Roots", Journal of Economic Surveys, Vol.4, No.3, pp.249-73.

Doan, T.A. (1996), RATS (Regression Analysis of Time Series) User's Manual Version 4, Estima: Evanston, IL 60201.

Doan, T.A., R.B. Litterman, and C.A. Sims (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions", Econometric Reviews, Vol.3, No.1, pp.1-144.

Doornik, J.A. and D.F. Hendry (1996), GiveWin - An Interface to Empirical Modelling, International Thomson Business Press, London.

Dua, P and S.C. Ray (1995), "A BVAR Model for the Connecticut Economy", Journal of Forecasting, Vol.14, No.3, pp.167-80.

Efron, B (1985), The Jackknife, the Bootstrap and Other Resampling Plans, J.W. Arrowsmith Ltd, Bristol.

Enders, W. (1995), Applied Econometric Time Series, John Wiley & Sons, New York.

—— (1996), RATS Handbook for Econometric Time Series, John Wiley & Sons, New York.

Engle, R.F. and C.W.J. Granger (1987), "Co-integration and Error Correction: Representation, Estimation, and Testing", Econometrica, Vol.55, No.2, pp.251-76.

Engle, R.F., C.W.J. Granger, S. Hylleberg and H.S. Lee (1993), "Seasonal Cointegration —— the Japanese Consumption Function", Journal of Econometrics, Vol.55, pp.275-98.

Engle, R.F. and D.F. Hendry (1993), "Testing Superexogeneity and Invariance in Regression Models", Journal of Econometrics, Vol.56, pp.119-39.

Engle, R.F., D.F. Hendry, and J.F. Richard (1983), "Exogeneity", Econometrica, Vol.51, No.2, pp.277-304.

Ericsson, N.R. (1988), "A Review of Data-Fit: An Interactive Econometric Modelling Package for IBM-Compatible PCs", Journal of Applied Econometrics, Vol.3, pp.319-32.

—— (1995), "Conditional and Structural Error Correction Models", Journal of Econometrics, Vol.69, pp.159-71.

Fair, R.C. and R.J. Shiller (1990), "Comparing Information in Forecasts from Econometric Models", the American Economic Review, Vol.80, No.3, pp.375-89.

Farmer, R.E.A. (1991), "The Lucas Critique, Policy Invariance and Multiple Equilibria", Review of Economic Studies, Vol.58, pp.321-32.

Favero, C. and D.F. Hendry (1992), "Testing the Lucas Critique: A Review", Econometric Reviews, 11(3), pp.265-306.

Figlewski, S. and P. Wachtel (1981), "The Formation of Inflationary Expectations", Review of Economics and Statistics, Vol.LXIII, pp.1-10.

Fildes, R. (1992), "The Evaluation of Extrapolative Forecasting Methods", International Journal of Forecasting, Vol.8, pp.81-98.

Fomby, T.B., F. George and Jr. Rhodes (1990), Co-integration, Spurious, and Unit Roots, JAI Press, Greenwich, Conn.

Fomby, T.B., R.C. Hill and S.R. Johnson (1984), Advanced Econometric Methods, Springer-Verlag, New York.

Freedman, D.A. (1986), "Reply", Journal of Business and Economic Statistics, Vol.4, pp.126-7.

Fu, Q. (1989), "The Choice of Optimal Locations of Overseas Agencies and Their Associated Cost-Benefit Analysis", Journal of Broadcasting Engineering, China's Radio & TV Publishers, No.36, pp.12-26.

Fuller, W.A. (1976), Introduction to Statistical Time series, Wiley, New York.

Gali, J. (1992), "How Well Does the IS-LM Model Fit Postwar U.S. Data", Quarterly Journal of Economics, May, pp.709-38.

Garcia-ferrer, A., R.A. Highfield, F. Palm, and A. Zellner (1987) "Macroeconomic Forecasting Using Pooled International Data", Journal of Business and Economic Statistics, Vol.5, No.1, pp.53-67.

George, E. (1997), "The 'Nervous' Governor", Business News, City University Business School, Summer, pp.5-8.

Geweke, J. (1982a), "Measurement of Linear Dependence and Feedback Between Multiple Time Series", Journal of the American Statistical Association, Vol.77, No.378, pp.304-13.

—— (1982b), "Rejoinder", Journal of the American Statistical Association, Vol.77, No.378, pp.323-4.

—— (1987), "Endogeneity and Exogeneity", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.134-6.

—— (1989), "Bayesian Inference in Econometric Models Using Monte Carlo Integration", Econometrica, Vol.57, No.6, pp.1317-39.

Godfrey, L.G. (ed) (1992), The Implementation and Constructive Use of Misspecification Tests in Econometrics, Manchester UP, Manchester.

Goel, P.K. and M.H. DeGroot (1981), "Information about Hyperparameters in Hierarchical Models", Journal of the American Statistical Association, Vol.76, pp.140-7.

Goel, P.K., A. Zellner and D.F. Bruno (1986), Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti, Amsterdam: Elsevier Science Publishers B.V.

Goldfeld, S.M. (1987), "Demand for Money: Empirical Studies", in Eatwell, J., M. Milgate and P. Newman (ed) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.770-5.

Goodhart, C.A.E. (1984), Monetary Theory and Practice: the UK Experience, Macmillan, London.

—— (1986), "Autobiographical Sketch", pp.322-3, in Blaug, M. (ed), Who's Who in Economics (2nd edn), MIT Press, Cambridge.

Gordon, R.J. (1976), "Can Econometric Policy Evaluations be Salvaged? —— A Comment", in Brummer, K. and Meltzer, A.H. (eds), The Phillips Curve and Labour Markets, Amsterdam: North-Holland, Carnegie - Rochester Conference Series on Public Policy, No.1, supplement to Journal of Monetary Economics, January, pp.47-61.

Gordon, S. (1997), "Stochastic Trends, Deterministic Trends, and Business Cycle Turning Points", Journal of Applied Econometrics,

Vol.12, pp.411-34.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", Econometrica, Vol.37, No.3, pp.424-38.

—— (1981), "Some Properties of Time Series Data and Their Use in Econometric Model Specification", Journal of Econometrics, Vol.16, pp.121-30.

—— (1986a), "Comment", Journal of Business and Economic Statistics, Vol.4, No.1, pp.16-7.

—— (1986b), "Developments in the Study of Cointegrated Economic Variables", Oxford Bulletin of Economics and Statistics, Vol.48, No.3, pp.213-28.

—— (1987), "Equilibrium, Causality and Error-Correction Models", Economic Notes, No.1, pp.5-21.

—— (1988), "Some Recent Developments in a Concept of Causality", Journal of Econometrics, Vol.39, pp.199-211.

—— (ed) (1991), Modelling Economic Series: Readings in Econometric Methodology, Oxford UP, Oxford.

Granger, C.W.J. and P. Newbold (1974), "Spurious Regressions in Econometrics", Journal of Econometrics, Vol.2, pp.111-20.

—— (1986), Forecasting Economic Time Series (2nd edn), Academic Press, London.

Graybill, F.A. (1969), Introduction to Matrices with Applications in Statistics, Wadsworth Publishing Company, Belmont, California.

Greene, W.H. (1997), Econometric Analysis (3rd edn), Upper Saddle River, N.J.

Gregory, A.W. and B.E. Hansen (1996), "Residual-Based Tests for Cointegration in Models with Regime Shifts", Journal of Econometrics, Vol.70, pp.99-126.

Griffiths, W.E., R.C. Hill and G.G. Judge (1993), Learning and Practicing Econometrics, Wiley, New York.

Hafer, R.W. and R.G. Sheehan (1989), "The Sensitivity of VAR Forecasts to Alternative Lag Structures", International Journal of Forecasting, Vol.5, pp.399-408.

Hall, S.G. (1989), "Practitioners Corner - Maximum Likelihood Estimation of Cointegration Vectors: An Example of the Johansen Procedure", Oxford Bulletin of Economics and Statistics, Vol.51, No.2, pp.213-18.

—— (1993), "Modelling Structural Change Using the Kalman Filter", Economics of Planning, Vol.26, No.1.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", Econometrica, Vol.57, pp.357-84.

Hannan, E.J. (1970), Multiple Time Series, John Wiley and Sons, London.

Hannan, E.J. and B.G. Quinn (1979), "The Determination of the Order of an Autoregression", Journal of the Royal Statistical Society, B41, pp.190-5.

Hansen, B. (1989), "Efficient Estimation of Cointegrating Vectors in the Presence of Deterministic Trends", Discussion Paper, University of Rochester.

Harvey, A.C. (1992), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge UP, Cambridge.

Hasan, M.A. (1993), "Policy Shocks and the Canadian Macroeconomy —— A Bayesian Vector Autoregression Approach", Economic Modelling, January, pp.81-8.

Haug, A.A. (1996), "Tests for Cointegration —— A Monte Carlo Comparison", Journal of Econometrics, Vol.71, pp.89-115.

Hendry, D.F. (1989), PC-Give: An Interactive Econometric Modelling System, Oxford UP, Oxford.

Hendry, D.F. and M.P. Clements (1992a), "On a Theory of Intercept Corrections in Macro-Econometric Forecasting", Discussion Paper, Institute of Economics and Statistics and Nuffield College, Oxford, pp.1-18.

—— (1992b), "Economic Forecasting", Discussion Paper, Institute of Economics and Statistics and Nuffield College, Oxford, pp.1-24.

Hendry, D.F. and J.A. Doornik (1996), Empirical Econometric Modelling Using PcGive 9.0 for Windows, International Thomson Business Press, London.

Hendry, D.F. and J.F. Richard (1982), "On the Formulation of Empirical Models in Dynamic Econometrics", Journal of Econometrics, Vol.20, pp.3-33.

—— (1983), "The Econometric Analysis of Economic Time Series", International Statistical Review, Vol.51, pp.111-63.

Hocking, R.R. (1976), "The Analysis and Selection of Variables in Linear Regression", A Biometrics Invited Paper, Biometrics, Vol.32, pp.1-49.

Holden, K. (1995), "Vector Autoregression Modeling and Forecasting", Journal of Forecasting, Vol.14, No.3, pp.159-66.

Holden, K. and D.A. Peel (1990), "On Testing for Unbiasedness and Efficiency of Forecasts", The Manchester School, Vol.LVIII,

pp.120-7.

Holland, P.W. (1986), "Statistics and Causal Inference", Journal of the American Statistical Association, Vol.81, No.396, pp.945-60.

Humphreys, P. (ed) (1994), Patrick Suppes: Scientific Philosopher, Vol.1, Probability and Probabilistic Causality, Kluwer, Dordrecht.

Hylleberg, S. and G.E. Mizon (1989), "Cointegration and Error Correction Mechanisms", Economic Journal, Vol.99 (Supplement), pp.113-25.

Jelenkowska, H. (1988), "Bayesian Estimation of the Variance Components in a General Random Linear Model", Bayesian Statistics, Vol.3, pp.653-6.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", Journal of Economic Dynamics and Control, Vol.12, pp.231-54.

—— (1989), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", Institute of Mathematical Statistics, University of Copenhagen, working paper.

—— (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", Econometrica, Vol.59, No.6, pp.1551-80.

—— (1992), "Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis", Journal of Econometrics, Vol.52, pp.389-402.

—— (1995), "Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration", Journal of Econometrics, Vol.69, pp.111-32.

—— (1996), Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford UP, Oxford.

Johansen, S. and K. Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration —— With Applications to the Demand for Money", Oxford Bulletin of Economics and Statistics, Vol.52, No.2, pp.169-210.

—— (1992), "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK", Journal of Econometrics, Vol.53, pp.211-44.

—— (1994), "Identification of the Long-Run and the Short-Run Structure —— An Application to the ISLM Model", Journal of Econometrics, Vol.63, pp.7-36.

Judge, G.G., R.C. Hill, W. Griffins, H. Lütkepohl and T.C. Lee (1982), Introduction to the Theory and Practice of Econometrics, J. Wiley & Sons, New York.

—— (1985), The Theory and Practice of Econometrics (2nd edn), J. Wiley & Sons, New York.

Judge, G.G., T.A. Yancey and M.E. Bock (1973), "Properties of Estimators after Preliminary Tests of Significance When Stochastic Restrictions Are Used in Regression", Journal of Econometrics, Vol.1, pp.29-48.

Kadane, J.B. (1984), Robustness of Bayesian Analyses, Elsevier Science Publishers B.V.

Kadane, J.B., N.H. Chan and L.J. Wolfson (1996), "Priors for Unit Root Models", Journal of Econometrics, Vol.75, pp.99-111.

Kadiyala, K.R. and S. Karlsson (1993), "Forecasting with Generalized Bayesian Vector Autoregressions", Journal of Forecasting, Vol.12, pp.365-78.

—— (1997), "Numerical Methods for Estimation and Inference in Bayesian VAR-Models", Journal of Applied Econometrics, Vol.12, pp.99-132.

Kakwani, N.C. (1968), "Note on the Unbiasedness of a Mixed Regression Estimator", Econometrica, Vol.36, No.3-4, pp.610-1.

Karisson, S. and H. Lütkepohl (1993), "Introduction to Multiple Time-Series", International Journal of Forecasting, Vol.9, No.4, pp.577-8.

Kazmier, L.J. (1988), Schaum's Outline of Theory and Problems of Business Statistics (2nd edn), McGraw-Hill, London.

Kennedy, P. (1992), A Guide to Econometrics (3rd edn), Basil Blackwell, Oxford.

King, R.G. and C.I. Plosser (1984), "Money, Credit, and Prices in a Real Business Cycle", American Economic Review, Vol.74, No.3, pp.363-80.

Kirchgässner, G. (1991), "Comments", European Economic Review, Vol.35, pp.918-22.

Koop, G. (1992), "'Objective' Bayesian Unit Root tests", Journal of Applied Econometrics, Vol.7, pp.65-82.

—— (1994), "Bayesian Semi-Nonparametric ARCH Models", Review of Economics and Statistics, Vol.LXXVI, pp.176-81.

Koop, G., J. Osiewalski and M.F.J. Steel (1995), "Bayesian Long-Run Prediction in Time Series Models", Journal of Econometrics, Vol.69, pp.61-80.

Koop, G. and S.M. Potter (1997), "Bayes Factors and Nonlinearity: Evidence From Economic Time Series", mimeo, University of Edinburgh and University of California, pp.1-35.

Kumar, V., R.P. Leone and J.N. Gaskins (1995), "Aggregate and Disaggregate Sector Forecasting Using Consumer Confidence Measures", International Journal of Forecasting, Vol.11, No.3, pp.361-77.

Laidler, D.E.W. (1985), The Demand for Money: Theories, Evidence, and Problems (3rd edn), Harper & Row, New York.

Leamer, E.E. (1982), "Sets of Posterior Means with Bounded Variance Priors", Econometrica, Vol.50, May, pp.725-36.

—— (1983), "Let's Take the Con out of Econometrics", American Economic Review, Vol.73, No.1 (March), pp.31-43.

—— (1991), "A Bayesian Analysis of the Determinants of Inflation", in Granger, C.W.J. (ed) (1991), Modelling Economic Series: Readings in Econometric Methodology, Oxford UP, Oxford.

Lee, E.O. (1987), "The Spurious Effects of Unit Roots on Vector Autoregressions", Journal of Econometrics, Vol.39, pp.251-66.

Lee, J.C., K.W. Lu and S.C. Horng (1992), "Technological Forecasting with Nonlinear Models", Journal of Forecasting, Vol.11, pp.195-206.

Lesage, J.P. (1992), "Scoring the Composite Leading Indicators: A Bayesian Turning Points Approach", Journal of Forecasting, Vol.11, pp.35-46.

Liang, K.Y. (1992) "On the Sign of the Optimal Combining Weights under the Error-Variance Minimizing Criterion", Journal of Forecasting, Vol.11, pp.719-23.

Lindley, D.V. (1987), "Bayes, Thomas (1702-1761)", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.207-8.

Lindley, D.V. and A.F.M. Smith (1972), "Bayes Estimates for the Linear Model", Journal of the Royal Statistical Society, Series B, Vol.34, pp.1-41.

Litterman, R.B. (1980), "A Bayesian Procedure for Forecasting with Vector Autoregression", mimeo, Massachusetts Institute of Technology.

—— (1986a), "Forecasting with Bayesian Vector Autoregressions —— Five Years of Experience", Journal of Business and Economic Statistics, Vol.4, No.1, pp.25-38.

—— (1986b), "Comment", Journal of Business and Economic Statistics, Vol.4, No,1, pp.17-24.

—— (1986c), "Specifying Vector Autoregressions for Macroeconomic Forecasting", in Goel, P. and A. Zellner (eds), Bayesian Inference and Decision Techniques, Elsever Science Publishers.

Litterman, R.B. and L. Weiss (1985), "Money, Real Interest Rates and

Output: a Reinterpretation of Postwar U.S. Data", Econometrica, Vol.53, pp.129-56.

Lucas, R.E. (1976a), "Econometric Policy Evaluation: A Critique", in Brummer, K. and Meltzer, A.H. (eds), The Phillips Curve and Labour Markets, Amsterdam: North-Holland, Carnegie - Rochester Conference Series on Public Policy, No.1, supplement to Journal of Monetary Economics, January, pp.19-46.

—— (1976b), "Understanding Business Cycles", pp.7-29.

Lucas, R.E. and T.J. Sargent (1981), Introduction, Rational Expectations and Econometric Practice, University of Minnesota Press, Minneapolis, Vol.1, pp.xi-xl.

Lütkepohl, H. (1991a), Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin.

—— (1991b), Forecasting Aggregated Vector ARMA Processes, Springer-Verlag, Berlin.

Machlup, F. (1958), "Equilibrium and Disequilibrium: Misplaced Concreteness and Disguised Politics", Economic Journal, March, reprinted in Machlup, F. (1963), Essays on Economic Semantics, Prentice-Hall Inc., Englewood Cliffs, N.J.

—— (1977), A History of Thought on Economic Integration, Macmillan, London.

MacKinnon, J.G. (1990), "Critical Values for Cointegration Tests", Discussion Paper, pp.1-14.

Maddala, G.S. (1988), Introduction to Econometrics (2nd edn), Macmillan, New York.

Mankiw, N.G. (1989), "Real Business Cycles: A New Keynessian Perspective", Journal of Economic Perspectives, Vol.3, No.3, pp.79-90.

Mariano, R.S. (1985), "Finite-Sample Properties of Stochastic Predictors in Nonlinear Systems: Some Initial Results", Warwick Economic Research Papers, 266, University of Warwick, pp.1-35.

Mariano, R.S. and Brown, B.W. (1983), "Asymptotic Behaviour of Predictors in a Non-Linear Simultaneous System", International Economic Review, Vol.21, pp.523-36.

—— (1991), "Stochastic-Simulation Tests of Nonlinear Econometric Models", Comparative Performance of U.S. Econometric Models, Oxford UP, New York, pp.250-9.

Marriott, J.M. (1988), "Reparameterisation for Bayesian Inference in ARMA Time Series", Bayesian Statistics, Vol.3, pp.701-4.

McCallum, B.T. (1980), "Rational Expectations and Macroeconomic Stabilization Policy: An Overview", Journal of Money, Credit, and Banking, Vol.12, No.4, pp.716-45.

Mckenzie, C.R. (1993), "Microfit 3.0: A Review - Software Reviews", Journal of Applied Econometrics, Vol.8, pp.413-9.

—— (1998), "Microfit 4.0 - Software Reviews", Journal of Applied Econometrics, Vol.13, pp.77-89.

McNees, S.K. (1986), "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts", Journal of Business and Economic Statistics, Vol.4, No.1, pp.5-15.

Mendenhall, W., Wackerly, D.D. and Scheaffer, R.L. (1990), Mathematical Statistics with Applications, Duxbury Press, Belmont, California.

Milgate, M. (1987), "Equilibrium: An Expectational Concept", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.177-83.

Mills, T.C. (1987), "Uncertainty in the U.K. Monetary Aggregates: Modelling Data Revisions in Economic Time Series", the Manchester School, pp.337-52.

—— (1989), "Four Paradoxes in U.K. GDP", Economics Letters, Vol.30, pp.287-90.

—— (1990a), Time Series Techniques For Economists, Cambridge UP Cambridge.

—— (1990b), "A Note on the Gibson Paradox during the Gold Standard", Prorations in Economic History, Vol.27, pp.277-86.

—— (1991a), "Equity Prices, Dividends and Gilt Yields in the UK: Cointegration, Error Correction and 'Confidence'", the Scottish Economic Society, pp.242-55.

—— (1991b), "Are Fluctuations in U.K. Output Transitory or Permanent?", the Manchester School, Vol.LIX, No.1, pp.1-11.

—— (1993), The Econometric Modelling of Financial Time Series, Cambridge UP, Cambridge.

Mills, T.C. and A.G. Mills (1991), "The International Transmission of Bond Market Movements", Bulletin of Economic Research, Vol.43, No.3, pp.273-81.

—— (1992), "Modelling the Seasonal Patterns in UK Macroeconomic Time Series", Journal of Royal Statistical Society, Vol.155, part 1, pp.61-75.

Mills, T.C., G.L. Pelloni and A. Zervoyianni (1993), "Sectoral Shifts and Unemployment Fluctuations: A test of the Lilien Hypothesis for the U.K.", Hull Economic Research Papers, No.204, pp.1-25.

Mills, T.C. and M.J. Stephenson (1985), "Forecasting Contemporaneous Aggregates and the Combination of Forecasts: the Case of the U.K. Monetary Aggregates", Journal of Forecasting, Vol.4 pp.273-81.

Mills, T.C. and M.P. Talor (1989), "Random Walk Components in Output and Exchange Rates: Some Robust Tests on UK Data", Bulletin of Economic Research, Vol.41, No.2, pp.123-33.

Mincer, J. and V. Zarnowitz (1969), "The Evaluation of Economic Forecasts", in Mincer J. (ed), Economic Forecasts and Expectations, National Bureau of Economic Research Studies in Business Cycles, No.19, pp.3-46, Columbia UP, New York.

Mizon, G.E. and J.F. Richard (1986), "The Encompassing Principle and Its Application to Testing Non-Nested Hypotheses", Econometrica, Vol.54, No.3, pp.657-78.

Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements", Econometrica, Vol.29, No.3, pp.315-35.

Nagar, A.L. and N.C. Kakwani (1964), "The Bias and Moment Matrix of a Mixed Regression Estimator", Econometrica, Vol.32, No.1-2, pp.174-82.

Nankervis, J.C. and N.E. Savin (1987), "Finite Sample Distributions of t and F Statistics in an AR(1) Model with an Exogenous Variable", Econometric Theory, Vol.3, pp.387-408.

Nelson, C.R. and C.I. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series —— Some Evidence and Implications", Journal of Monetary Economics, 10(2), pp.139-62.

Ohanian, L.E. (1988), "The Spurious Effects of Unit Roots on Vector Autoregressions : A Monte Carlo Study", Journal of Econometrics, Vol.39, pp.251-66.

—— (1990), "A Note on Spurious Inference in a Linearly Detrended Vector Autoregression", Review of Economics and Statistics, Vol.73, pp.568-71.

Oliver, R.M. (1987), "Bayesian Forecasting With Stable Seasonal Patterns", Journal of Business and Economic Statistics, Vol.5, No.1, pp.77-85.

Osborn, D.R. and J.P. Smith (1989), "The Performance of Periodic Autoregressive Models in Forecasting Seasonal U.K. Comparison", Journal of Business and Economic Statistics, Vol.7, No.1, pp.117-27.

Osterwald-Lenum, M. (1993), "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Tests Statistics", Oxford Bulletin of Economics and Statistics, Vol.54, pp.461-72.

Owen, D. (1986), Money, Wealth and Expenditure: Integrated Modelling of Consumption and Portfolio Behaviour, Cambridge UP, Cambridge.

Pagan, A.R. and M.R. Wickens (1989), "A Survey of Some Recent Econometric Methods", Economic Journal, Vol.99, pp.962-1025.

Pain, N. and A. Britton (1992), "National Institute Economic Forecasts 1986 to 1991: Some Tests of Forecast Properties", National Institute Economic Review, August, pp.81-93.

Palm, F.C. (1995), "Bayesian Model Selection and Prediction with Empirical Applications: Comments", Journal of Econometrics, Vol.69, pp.333-5.

Palm, F.C. and A. Zellner (1992), "To Combine or not to Combine? Issues of Combining Forecasts", Journal of Forecasting, Vol.11, pp.687-701.

Parzen, E. (1982), "Comment", Journal of the American Statistical Association, Vol.77, No.378, pp.320-2.

Paulus, J.D. (1975), "Mixed Estimation of a Complete System of Consumer Demand Equations", Annals of Economic and Social Measurement, Vol.4, No.1, pp.117-31.

Perron, P. (1988), "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach", Journal of Economic Dynamics and Control, Vol.12, pp.297-332.

—— (1989), "Testing for a Unit Root in a Time Series with a Changing Mean", Journal of Business and Economic Statistics, Vol.8, pp.153-62.

Pesaran, M.H. (1987), "Econometrics", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.8-20.

Pesaran, M.H. and B. Pesaran (1991), Microfit 3.0 —— An Interactive Econometric Software Package, Oxford UP, Oxford.

—— (1997), Working with Microfit 4.0 —— Interactive Econometric Analysis, Oxford UP, Oxford.

Phillips, P.C.B. (1986), "Understanding Spurious Regressions in Econometrics", Journal of Econometrics, Vol.33, pp.311-40.

—— (1995a), "Bayesian Model Selection and Prediction with Empirical Applications", Journal of Econometrics, Vol.69, pp.289-331.

—— (1995b), "Bayesian Prediction —— A Response", Journal of Econometrics, Vol.69, pp.351-65.

—— (1995c), "Nonstationary Time Series and Cointegration", Journal of Applied Econometrics, Vol.10, pp.87-94.

Phillips, P.C.B. and S.N. Durlauf (1986), "Multiple Time Series Regression with Integrated Processes", Review of Economic Studies, Vol.LIII, pp.473-95.

Pierce, D.A. (1982), "Comment", Journal of the American Statistical Association, Vol.77, No.378, pp.315-6.

Pindyck, R.S. and D.L. Rubinfeld (1991), Econometric Models and Economic Forecasts (3rd edn), McGraw-Hill, New York.

Plosser, C.I. (1989), "Understanding Real Business Cycles", Journal of Economic Perspective, Vol.3, No.3, pp.51-77.

Poirier, D.J. (1988), "Bayesian Diagnostic Testing in the General Linear Normal Regression Model", Bayesian Statistics, Vol.3, pp.725-32.

Poole, W. (1967), "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model", Quarterly Journal of Economics, Vol.LXXXIV, pp.197-216.

Radaelli, G. (1992) "Exchange Rate Models: Monetary or Portfolio Balance Effects?", Discussion Paper, Lehman Brothers and Birkbeck College, pp.1-22.

Rao, C.R. (1973), Linear Statistical Inference and Its Applications (2nd edn), Wiley, New York.

Reinsel, G.C. and S.K. Ahn (1988), "Asymptotic Distribution of the Likelihood Ratio Test for Cointegration in the Nonstationary Vector AR Model", Technical Report, Department of Statistics, University of Wisconsin, Madison.

Richard, J.F. (1995), "Bayesian Model Selection and Prediction with Empirical Applications Discussion", Journal of Econometrics, Vol.69, pp.337-49.

Richardson, P. (1992), "Simulating the European Economies under Alternative Monetary Policy Assumptions", Journal of Forecasting, Vol.11, pp.389-421.

Robert, F.E. and D.F. Hendry (1993), "Testing Superexogeneity and Invariance in Regression Models", Journal of Econometrics, Vol.56, pp.119-39.

Robert, E. and J. Lucas (1976), "Econometric Policy Evaluation: A Critique", in Brunner-Meltzer, The Phillips Curve & Labour Markets, Canngie-Rochester Conferences, Vol.I, JME Supplement.

—— (1976), "Understanding Business Cycles", University of Chicago.

Robinson, P.M. (1989), "Hypothesis Testing in Semiparametric and Nonparametric Models for Econometric Time Series", Review of Economic Studies, Vol.56, pp.511-34.

Rothenberg, T.J. (1973), Efficient Estimation with a Priori Information, Yale UP, New Haven (Conn).

Runkle, D.E. (1987a), "Vector Autoregressions and Reality", Journal

of Business & Economic Statistics, Vol.5, No.4, pp.437-42.

—— (1987b), "Reply", Journal of Business & Economic Statistics, Vol.5, No.4, p.454.

Rust, J. (1988), "Comment on Poirier: The Subjective Perspective of a 'Spiritual Bayesian'", Journal of Economic Perspectives, Vol.2, No.1, pp.145-51.

Sargan, J.D. (1964), "Wages and Prices in the United Kingdom: A Study in Econometric Methodology", Econometric Analysis for National Economic Planning, ed. by P.E. Hart *et al.* Butterworths, London.

Sargent, T.J. (1984), "Autoregressions, Expectations, and Advice", American Economic Reviews, Vol.74, No.2, pp.408-16.

—— (1987), "Rational Expectations", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.76-9.

Schwarz, G. (1978), "Estimating the Dimension of a Model", Annals of Statistics, Vol.6, pp.461-4.

Shoesmith, G.L. (1990), "The Forecasting Accuracy of Regional Bayesian VAR Models with Alternative National Variable Choices", International Regional Science Review, Vol.13, No.3, pp.257-69.

—— (1992), "Co-Integration, Error Correction and Improved Medium-Term Regional VAR Forecasting", Journal of Forecasting, Vol.11, pp.91-109.

Simon, H.A. (1987), "Causal Inference", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.380-3.

Sims, C.A. (1980a), "Macroeconomics and Reality", Econometrica, Vol.48, No.1, pp.1-48.

—— (1980b), "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered", American Economic Association, Vol.70, No.2, pp.250-7.

—— (1982), "Policy Analysis with Econometric Models", Brookings Papers on Economic Activity, pp.107-64.

—— (1987), "Comment", Journal of Business & Economic Statistics, Vol.5, No.4, pp.443-49.

—— (1989), "A Nine Variable Probabilistic Macroeconomic Forecasting Model", Discussion Paper, No.14, Institute for Empirical Macroeconomics, University of Minnesota.

—— (1991), "Comments", European Economic Review, Vol.35, pp.922-32.

—— (1992), "Interpreting the Macroeconomic Time Series Facts ——

the Effects of Monetary Policy", European Economic Review, Vol.36, pp.975-1011.

Sims, C.A., J.H. Stock and M.W. Watson (1990), "Inference in Linear Time Series Models with Some Unit Roots", Econometrica, Vol.58, pp.113-44.

Smith, G. (1980), "Further Evidence on the Value of a Priori Information", Journal of Finance, Vol.XXXV, No.1, pp.181-9.

Smith, G. and W. Brainard (1976), "The Value of a Priori Information in Estimating a Financial Model", Journal of Finance, Vol.XXXI, No.5, pp.1299-322.

Smith, T.M.F. (1988), "To Weight or not to Weight, That is the Question", Bayesian Statistics, Vol.3, pp.437-51.

Söderlind, P. and A. Vredin (1996), "Applied Cointegration Analysis in the Mirror of Macroeconomic Theory", Journal of Applied Econometrics, Vol.11, pp.363-81.

Spencer, D.E. (1993), "Developing a Bayesian Vector Autoregression Forecasting Model", International Journal of Forecasting, Vol.9, pp.407-21.

Stewart, J. (1986), Understanding Econometrics (2nd edn), Hutchinson, London.

—— (1991), Econometrics, Philip Allan, New York.

Stewart, M.B. and K.F. Wallis (1990), Introductory Econometrics (2nd edn), Basil Blackwell, Oxford.

Stock, J.H. and M.W. Watson (1988a), "Variable Trends in Economic Time Series", Journal of Economic Perspectives, Vol.2, pp.147-74.

—— (1988b), "Testing for Common Trends", Journal of the American Statistical Association, Vol.83, pp.1097-107.

Stockman, A.C. (1990), "International Transmission and Real Business Cycle Models", New Classical Macroeconomics in the Open Economy, Vol.80, No.2, pp.134-8.

Swamy, P.A.V.B. and J.S. Mehta (1969), "On Theil's Mixed Regression Estimator", Journal of the American Statistical Association, Vol.64, pp.273-6.

Theil, H. (1963), "On the Use of Incomplete Prior Information in Regression Analysis", Journal of the American Statistical Association, Vol.58, pp.401-14.

—— (1971), Principles of Econometrics, Wiley, New York.

—— (1974a), "Mixed Estimation Based on Quasi-Prior Judgments", European Economic Review, Vol.5, pp.33-40.

—— (1974b), Introduction to Econometrics, Prentice-Hall, Englewood Cliffs, N.J.

Theil, H. and A.S. Goldberger (1961), "On Pure and Mixed Statistical Estimation in Economics", International Economic Review, Vol.2, No.1, pp.65-78.

Timmermann, A. (1995), "Cointegration Tests of Present Value Models with a Time-Varying Discount Factor", Journal of Applied Econometrics, Vol.10, pp.17-31.

Tjøstheim, D. (1981), "Granger-Causality in Multiple Time Series", Journal of Econometrics, Vol.17, pp.156-76.

Todd, R.M. (1984), "Improving Economic Forecasting with Bayesian Vector Autoregression", Quarterly Review of the Federal Reserve Bank of Minneapolis, Fall, pp.18-29.

Trader, R. (1983), "A Bayesian Technique for Selecting a Linear Forecasting Model", Management Science, Vol.29, No.5, pp.622-32.

Tremayne, A.R. (1992), "Integrated Variables and Tests for Unit Roots", in Godfrey, L.G. (ed) (1992), The Implementation and Constructive Use of Misspecification Tests in Econometrics, Manchester UP, Manchester.

Tremayne, A.R. and N. Davids (1994), "Review of Statgraphics", Journal of Applied Econometrics, Vol.9, pp.335-41.

Tremayne, A.R. and B.P.M. McCabe (1993), Elements of Modern Asymptotic Theory with Statistical Applications, Manchester UP, Manchester.

—— (1995), "Testing a Time Series for Difference Stationarity", Annals of Statistics, Vol.23, No.3, pp.1015-28.

Tremayne, A.R. and P.C. Smith (1992), "On the Information Content of Time Series of Accounting Earnings", Journal of Business Finance and Accounting, 19(3), pp.363-85.

Trivedi, P.K. (1980), "Small Samples and Collateral Information: An Application of the Hyperparameter Model", Journal of Econometrics, Vol.12, pp.301-18.

Urbain, J.P. (1993), Exogeneity in Error Correcting Models, Springer-Verlag, Berlin.

—— (1995), "Partial Versus Full System Modelling of Cointegrated Systems - An Empirical Illustration", Journal of Econometrics, Vol.69, pp.177-210.

Vahid, F. and R.F. Engle (1993), "Common Trends and Common Cycles", Journal of Applied Econometrics, Vol.8, pp.341-60.

Wallis, K.F. (1986), Models of the UK Economy, Oxford UP, Oxford.

—— (1989), "Macroeconomic Forecasting: A Survey", Economic

Journal, Vol.99, pp.28-61.

—— (1994), Macroeconometric Modelling, E. Elgar, Aldershot.

Wallis, K.F., P.G. Fisher, J.A. Longbottom, D.S. Turner and J.D. Whitley (1987), Models of the UK Economy, Oxford UP, Oxford.

Wei, W.W.S. (1982), "Comment: The Effects of Systematic Sampling and Temporal Aggregation on Causality —— A Cautionary Note", J.A.S.A. Vol.77, No.378, pp.316-9.

West, K.D. (1987), "A Specification Test for Speculative Bubbles", Quarterly Journal of Economics, Vol.102, pp.553-80.

—— (1988), "Asymptotic Normality, When Regressors have a Unit Root", Econometrica, Vol.56, pp.1397-418.

West, M. and J. Harrison (1989), Bayesian Forecasting and Dynamic Models, Springer-Verlag, New York.

Westlund, A.H. (1993), "Business-Cycle Forecasting", Journal of Forecasting, Vol.12, No.3-4, pp.187-96.

Whitley. J.D. (1991), "Comparative Simulation Analysis of the European Multi-Country Models", Journal of Forecasting, Vol.11, pp.423-58.

Wickens, M.R. (1996), "Interpreting Cointegrating Vectors and Common Stochastic Trends", Journal of Econometrics, Vol.74, pp.255-71.

Wold, H. (1938), A Study in the Analysis of Stationary Time-Series, Almqvist and Wiksells, Uppsala.

Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, John Wiley & Sons, New York.

—— (1982), "Comment", Journal of the American Statistical Association, Vol.77, No.378, pp.313-4.

—— (1985), "Bayesian Econometrics", Econometrica, Vol.53, No.2, pp.253-67.

—— (1986), "Biased Predictors, Rationality and the Evaluation of Forecasts", Economics Letters, Vol.21, pp.45-8.

—— (1987), "Bayesian Inference", in Eatwell, J., M. Milgate and P. Newman (edi) (1987), The New Palgrave —— A Dictionary of Economics, Macmillan Press Ltd, London, pp.208-18.

Zellner, A. and C. Hong (1989), "Forecasting International Growth Rates Using Bayesian Shrinkage and Other Procedures", Journal of Econometrics, Vol.40, pp.183-202.