Journal of Forecasting, Vol. 14, 159–166 (1995)

Vector Autoregression Modelling and Forecasting

KEN HOLDEN
The Business School, Liverpool John Moores University, U.K.

ABSTRACT

This paper provides an introduction to vector autoregression models, explaining their origins and their use for modelling and forecasting. The recent developments of structural modelling and the treatment of non-stationary variables are also considered.

KEY WORDS economic forecasting; vector autoregression; structural VAR; non-stationary variables

INTRODUCTION

Vector autoregression (VAR) models are, in principle, simple multivariate models in which each variable is explained by its own past values and the past values of all the other variables in the system. They have been used extensively for applied modelling and forecasting since the work of Sims (1980). To set the scene, a simple VAR model to explain the two variables x and y is given by:

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 y_{t-1} + \alpha_4 y_{t-2} + \varepsilon_{1t} \quad (1)$$

$$y_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 y_{t-1} + \beta_4 y_{t-2} + \varepsilon_{2t} \quad (2)$$

where the ε_i are random errors which, in general, will be correlated, and, for exposition purposes, just two lagged values of each variable are included in the equations. Notice that the right-hand-side variables are the same in each equation so that ordinary least squares is a valid estimation method. Also note that in this example, with two variables, a maximum lag of two periods and two constants, there are ten unknown parameters to be estimated in equations (1) and (2), and that, more generally, a system of k variables with l lags on each variable and a constant in each equation would have $(k^2 l) + k$ parameters. The general VAR model can be written

$$\mathbf{Y}_t = \mathbf{A}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (3)$$

where $\mathbf{A}$ is a matrix of polynomials in the lag operator, $\mathbf{Y}$ is the vector of variables, and $\boldsymbol{\varepsilon}$ is a vector of random errors.

THE STATISTICAL BACKGROUND

VAR models arise from two separate strands of research. In the time-series analysis literature, the pioneering work of Wold (1938, Theorem 7) proved that any stationary variable can be

CCC 0277–6693/95/030159–08 *Received October 1994*

uniquely represented by a (possibly infinite) moving average (MA) process. That is, ignoring any deterministic component which can be removed by subtraction,

$$\begin{aligned} y_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \\ &= (1 + \theta_1 L + \theta_2 L^2 + \cdots)\varepsilon_t \\ &= \theta(L)\varepsilon_\tau \end{aligned} \tag{4}$$

where L is the lag operator and ε is a white-noise error (i.e. with a mean of zero, a constant variance, and zero covariances). Here, y is expressed in terms of current and past random shocks. Also, Whittle (1963, p. 21) proposed the alternative of a (possibly infinite) autoregressive (AR) representation for a stationary variable. Again ignoring any deterministic component,

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \cdots + \varepsilon_t \tag{5}$$

so that y is expressed in terms of its own past values and a white-noise error. Now equation (5) can be written

$$\begin{aligned} \varepsilon_t &= (1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 \ldots) y_t \\ &= \phi(L) y_t \end{aligned} \tag{6}$$

The representations in equations (4) and (5) can involve a possibly infinite number of parameters but they provide the groundwork for the univariate autoregressive moving-average (ARMA) models of Box and Jenkins (1970). These can be written

$$\phi(L) y_t = \theta(L)\varepsilon_t \tag{7}$$

where both $\phi(L)$ and $\theta(L)$ are low-order polynomials in the lag operator, L. While these models have been widely adopted for forecasting, they are univariate and a natural extension is to include one or more exogenous variables (to give the ARMAX model). However, for the such a model to be used for forecasting requires forecasts of future values of these exogenous variables. As an alternative, vector ARMA models (Tiao and Box, 1981) have been suggested, in which all variables are endogenous. For a vector, $\mathbf{Y}$, of stationary variables, these can be represented as

$$\mathbf{B}\mathbf{Y}_t = \mathbf{C}\boldsymbol{\varepsilon}_t \tag{8}$$

where $\mathbf{B}$ and $\mathbf{C}$ are matrix-valued polynomials in the lag operator and $\boldsymbol{\varepsilon}$ is a multivariate white-noise process. In the special case where $\mathbf{B}$ is invertible, equation (8) can be written

$$\mathbf{Y}_t = \mathbf{B}^{-1}\mathbf{C}\boldsymbol{\varepsilon}_t$$

which is the MA representation and when $\mathbf{C}$ is invertible,

$$\mathbf{C}^{-1}\mathbf{B}\mathbf{Y}_t = \boldsymbol{\varepsilon}_t \qquad \text{or} \qquad \mathbf{D}\mathbf{Y}_t = \boldsymbol{\varepsilon}_t \tag{9}$$

which is the AR representation. Now $\mathbf{D}$ includes constants and polynomials in the lag operator and so can be written

$$\mathbf{D} = \mathbf{F} + L\mathbf{G}$$

where $\mathbf{F}$ is a matrix of constants and $\mathbf{G}$ includes constants and terms in the lag operator. Substituting into equation (9) gives

$$\mathbf{F}\mathbf{Y}_t = -\mathbf{G}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t$$

and

$$\mathbf{Y}_t = -\mathbf{F}^{-1}\mathbf{G}\mathbf{Y}_{t-1} + \mathbf{F}^{-1}\boldsymbol{\varepsilon}_t \tag{10}$$

giving the same form as equation (3), the general VAR model, except for the inclusion of $\mathbf{F}^{-1}$ in the final term which occurs because of the possibility of the error terms being contemporaneously correlated.

THE ECONOMETRIC BACKGROUND

In the econometrics literature, while the main stimulus for much recent work on VAR models is the paper by Sims (1980), the basic idea of using an unrestricted vector of past values of variables for forecasting is implicit in the use of a recursive system by Wold (1938) to represent a causal chain, which links with the work of Tinbergen (1939) on forecasting using an economic model. Following Zellner and Palm (1974), a general dynamic linear simultaneous econometric model can be written in the structural form

$$\mathbf{H}\mathbf{Y}_t + \mathbf{J}\mathbf{X}_t = \mathbf{K}\boldsymbol{\varepsilon}_{1t} \tag{11}$$

where $\mathbf{Y}$ and $\mathbf{X}$ are vectors of endogenous and exogenous variables, $\mathbf{H}$, $\mathbf{J}$, and $\mathbf{K}$ are matrix-valued polynomials in the lag operator, and $\boldsymbol{\varepsilon}_1$ is a vector of white-noise random errors with $\mathrm{E}(\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{1t}') = \mathbf{I}$. In standard economic modelling prior restrictions are imposed on the $\mathbf{H}$, $\mathbf{J}$ and $\mathbf{K}$ matrices, the coefficients are estimated and forecasts can then be made if future values of $\mathbf{X}$ are available. However, if we ignore the exogenous variables, by assuming that they can be included with the endogenous variables and put

$$\mathbf{H} = \mathbf{H}_1 + L\mathbf{H}_2$$

where $\mathbf{H}_1$ is a matrix of constants which gives the interrelationships between the current values of the variables, and $\mathbf{H}_2$ is a matrix of constants and polynomials in the lag operator; then (11) becomes

$$\mathbf{H}_1\mathbf{Y}_t + \mathbf{H}_2\mathbf{Y}_{t-1} = \mathbf{K}\boldsymbol{\varepsilon}_{1t}$$

and the reduced form is

$$\mathbf{Y}_t = -\mathbf{H}_1^{-1}\mathbf{H}_2\mathbf{Y}_{t-1} + \mathbf{H}_1^{-1}\mathbf{K}\boldsymbol{\varepsilon}_{1t} \tag{12}$$

This can be written

$$\mathbf{Y}_t = \mathbf{A}^*\mathbf{Y}_{t-1} + \boldsymbol{e}_t \tag{13}$$

which is the same form as equation (3). Notice that the error term in equation (13) has

$$\begin{aligned}\mathrm{E}(\boldsymbol{e}_t\boldsymbol{e}_t') &= \mathrm{E}[\mathbf{H}_1^{-1}\mathbf{K}\boldsymbol{\varepsilon}_{1t}(\mathbf{H}_1^{-1}\mathbf{K}\boldsymbol{\varepsilon}_{1t})'] \\ &= \mathrm{E}[\mathbf{H}_1^{-1}\mathbf{K}\mathbf{K}'(\mathbf{H}_1^{-1})'] = \boldsymbol{\Omega}\end{aligned} \tag{14}$$

say, which is not the unit matrix, so that care must be taken in making inferences about the behaviour of the model in response to shocks. This is considered in the next section.

Here it has been assumed that the structural model is linear but the VAR model can also be considered as a linear approximation to the reduced form of any non-linear structural model.

Each of these approaches to VAR models starts from the statistical theory. In contrast, Sims

(1980) starts from conventional economic modelling procedures and questions the validity of restrictions from economic theory (such as imposing zeros in the **H**, **J**, and **K** matrices in equation (11) above), the arbitrary normalization of equations, the imposition of lag patterns, and the assumptions concerning identification. Sims rejects conventional economic models in favour of an atheoretic reduced-form approach based on VAR models. The only assumptions required in this approach are that the modeller has an accurate list of relevant variables and knows the form (i.e. levels or differences) in which they are to appear. The VAR model can then be specified and the maximum lag length can be determined empirically by a simple test (see Sims, 1980, p.17) after the estimation of equation (3) by ordinary least squares. But, as Robertson and Wickens (1994) point out, even deciding to use real rather than nominal variables means that the modeller's judgement is needed.

The attraction of Sims's methodology is that the VAR model is viewed as an unrestricted reduced form from a structural model. If, as Sims argues, the restrictions usually imposed on economic models from economic theory are invalid, it is better to ignore them and let the data determine the model. However, should those restrictions be valid, there will be a loss of efficiency in ignoring them.

ESTIMATION PROBLEMS

One obvious problem with the general VAR model (3) is the large number of parameters to be estimated. For example, in Sims (1980) the models have six variables and the lag length is initially eight, giving in each equation 48 coefficients on the lagged variables to be estimated, plus the constant term, while there are 108 quarterly observations for the USA and 96 for Germany. Apart from the multicollinearity between the different lagged variables leading to imprecise coefficient estimates, the large number of parameters leads to a good within-sample fit but poor forecasting accuracy because, according to Litterman (1986, p.2), 'parameters fit not only the systematic relationships ... but also the random variation'. Instead, Litterman proposes a Bayesian vector autoregression (BVAR) model (Doan *et al.*, 1984) in which the prior assumption (the 'Minnesota' or 'Litterman' prior) is that each variable can be modelled by a random walk with drift

$$x_t = x_{t-1} + c + \varepsilon_t \tag{15}$$

where c is a constant and ε is a white-noise disturbance. All other parameters are assumed to have a mean of zero but with a non-zero variance which is larger on lags of the dependent variable than on other variables. Also, the parameters are assumed to be uncorrelated and to have variances which decline as the lag length increases. The effect of these assumptions is to allow the data to determine the pattern of the coefficients, with it being easier for lagged values of the dependent variable to enter the equations. Kadiyala and Karlsson (1993) discuss the merits of alternatives to the Minnesota prior. The evidence is that the forecasts produced by BVAR models are at least as accurate as forecasts from traditional economic models (see McNees, 1986; Dua and Ray, 1995; Curry *et al.*, 1995; Dua and Smyth, 1995; Sarantis and Stewart, 1995).

An alternative to the BVAR method of reducing the number of parameters estimated is proposed by Gilbert (1995). This involves a model-reduction procedure whereby each smaller model retains the dynamic characteristics of the initial model.

INNOVATION ANALYSIS

Sims summarizes his estimated VAR model by examining innovations, that is, the response of the system to typical random shocks. The VAR model (3) can be written

$$(\mathbf{I} - \mathbf{A}L)\mathbf{Y}_t = \boldsymbol{\varepsilon}_t$$

and the MA representation is

$$\mathbf{Y}_t = (\mathbf{I} - \mathbf{A}L)^{-1}\boldsymbol{\varepsilon}_t = \boldsymbol{\Theta}(L)\boldsymbol{\varepsilon}_t$$

which gives the dynamic response to shocks. However, there is a problem in the interpretation of these innovations. From comparing equations (12) and (13) the relationship between the structural disturbances ($\boldsymbol{\varepsilon}_1$) and the reduced-form disturbances ($\mathbf{e}$) is

$$\mathbf{e}_t = \mathbf{H}_1^{-1}\mathbf{K}\boldsymbol{\varepsilon}_{1t} \tag{16}$$

and estimation of the reduced form does not allow the parameters in $\mathbf{H}_1$ and $\mathbf{K}$ to be identified.

As an illustration, consider the simple VAR model (1) and (2) where the error terms are contemporaneously correlated, so that a shock to equation (1) also has a current-period effect on equation (2), and vice versa. We can write

$$\varepsilon_{1t} = \gamma\varepsilon_{yt} + \varepsilon_{xt} \tag{17}$$

$$\varepsilon_{2t} = \delta\varepsilon_{xt} + \varepsilon_{yt} \tag{18}$$

where the error terms have been split into ε_{xt}, a component due to an innovation in x, and ε_{yt}, a component due to an innovation in y. The parameters γ and δ measure the importance of the effect of a shock to y on x and a shock to x on y. From estimating equations (17) and (18) it is not possible to separate out the effects of the separate shocks, since both shocks affect x and y.

However, this changes if the restriction $\gamma = 0$ is imposed, and a current-period shock to y does not affect x:

$$\varepsilon_{1t} = \varepsilon_{xt} \tag{19}$$

while $\delta \neq 0$, so that the shocks to both x and y affect y. In this case, equation (19) gives the effect of a shock to x on x, and (18) gives the combined effect of the shocks to x and y on y. Since (18) can be estimated from a knowledge of ε_1 and ε_2, an estimate of δ can be found as the ratio of the covariance of ε_{1t} and ε_{2t} to the variance of ε_{1t}. That is, by imposing a restriction on (17) it is possible to determine the separate effects of shocks. The restriction here is known as a 'causal ordering' because the pattern is that first x is affected by a shock ε_{xt} and then y is affected by both ε_{xt} and ε_{yt} so that contemporaneous causality runs from x to y.

In this simple example the only other causal ordering occurs when $\gamma \neq 0$ and $\delta = 0$, so that a shock to y affects both y and x, while a shock to x affects only x.

Sims's solution to the identification problem is to order the variables from the most pervasive—a shock to this variable affects all the other variables in the current period —to least pervasive—a shock to this variable does not affect any other variable in the current period. The result is a triangular pattern in the residuals of the VAR equations, with one residual included in the first equation, two in the second, and so on. In practice, the VAR model is estimated by ordinary least squares and the covariance matrix of the residuals, $\boldsymbol{\Omega}$, is estimated. The Choleski decomposition of this matrix is performed to give the lower triangular matrix, $\boldsymbol{\lambda}$, say, which satisfies

$$\boldsymbol{\Omega} = \boldsymbol{\lambda}\boldsymbol{\lambda}'$$

and by comparison with equation (14),

$$\boldsymbol{\lambda} = \mathbf{H}_1^{-1}\mathbf{K}$$

This triangularization imposes restrictions which allow the effects of the shocks to be identified and so they can be interpreted as structural rather than reduced-form shocks. However, in a system of k equations there are $k!$ possible orderings and Cooley and LeRoy (1985) criticize the procedure as being arbitrary since economic theory is not clear on the choice of causal ordering.

To overcome this problem, 'structural' VAR models have been proposed. One approach, suggested by Bernanke (1986) and Sims (1986), is to use economic theory to determine what contemporaneous structural restrictions should be imposed on the $\mathbf{H}_1$ and $\mathbf{K}$ matrices. Alternatively, Shapiro and Watson (1988) and Blanchard and Quah (1989) exploit long-run properties to impose identifying restrictions on the parameters. An excellent review of these models is provided by Keating (1992). In this issue, Ioannidis *et al.* (1995) provide an example of imposing restrictions on the structural parameters by assuming that world output has no feedback from UK or Welsh economic variables, and that UK variables have no feedback from Welsh economic variables.

COINTEGRATION, ERROR CORRECTION, AND VAR MODELS

One area of controversy since Sims (1980) is whether the variables included in a VAR should be stationary or not. Sims includes the levels of such non-stationary variables as money and prices while other researchers (for example, Lupoletti and Webb, 1986) transform the variables to be stationary by forming rates of change. However, the development of tests for stationarity and the interest in cointegration following the work of Engle and Granger (1987) and Engle and Yoo (1987) has clarified the situation and Robertson and Wickens (1994) provide an up-to-date review.

If all the variables under consideration are stationary (which can be checked by the Dickey and Fuller, 1981, test) then the VAR should be estimated with these variables. Any shocks to stationary variables can have only a temporary effect.

When the variables are not stationary the situation is rather more complicated since the procedure depends on whether the variables are cointegrated. If they are not cointegrated, the correct approach is to transform the variables to become stationary (usually by first-differencing them) and then estimate the VAR in the usual way. But care must be taken in interpreting shocks because for first-differenced variables a shock will have a temporary effect on the change of the variable and a permanent effect on its level.

For non-stationary variables which are cointegrated the use of only first-differenced variables in the VAR is incorrect. This is because the Granger representation theorem (see Engle and Granger, 1987, p.255) states that the cointegrated variables are related through an error-correction model (ECM) which includes the differenced variables and also the levels of the cointegrated variables. Therefore just using differenced variables omits important terms and will, in general, lead to poor forecasts. The correct method is to estimate the vector ECM, which is a VAR with the addition of a vector of cointegrating residuals. The alternative of estimating a VAR in levels of the variables avoids these problems but, by ignoring the restrictions from cointegration, produces inefficient forecasts. These issues are discussed by Webb (1995) in the context of forecasting inflation.

When there is a mixture of stationary and non-stationary variables the situation is more complex, but Sims *et al.* (1990) argue that the Bayesian approach to estimating VARs is still

valid and no special account of the non-stationarity need be taken. The benefits of using Bayesian methods to estimate vector ECM models are explored by Joutz *et al.* (1995) and, in the context of non-linear dynamics, by Pinder and Shoesmith (1995).

CONCLUSIONS

VAR modelling has made much progress since Sims (1980). It is now generally accepted that the Bayesian approach to estimation results in better forecasts than the unrestricted VAR. The question of the interpretation of the innovations analysis is more problematical. Sims's rejection of economic theory as a guide to identification of the structural equations, linked with his use of it in determining the causal orderings of variables, is still controversial. Some use of economic theory is necessary, even if only in deciding which variables should be selected. The argument is over how much should be used. While structural VARs avoid the arbitrariness of the Choleski decomposition this is at the cost of imposing a particular economic structure. The development of vector ECM's, which exploit the long-term properties of the data, offers the promise of better forecasts but whether this will generally yield substantial improvements in short-term forecasting accuracy remains to be seen.

ACKNOWLEDGEMENTS

I am grateful to the following who acted as referees for this issue: Isabel Andrade, Peter Baron, David Byers, Keith Cuthbertson, Philip Davis, Pami Dua, Javier Fernandez-Macho, Anita Ghatak, Stephen Hall, Mohammed Hasan, Darryl Holden, Chris Ioannides, Jason Laws, Ronald MacDonald, Kent Matthews, Bob McClelland, Bruce McCullough, Terence Mills, Kerry Patterson, David Peel, Bahram Pesaran, Zacharias Psaradakis, Nicholas Sarantis, Gary Shoesmith, Haiyan Song, John Thompson, Robert Trost and Ruey Tsay.

REFERENCES

Bernanke, B. S., 'Alternative explanations of the money–income correlation', *Carnegie-Rochester Conference Series on Public Policy* (1986) 49–100.

Blanchard, O. J. and Quah, D. 'The dynamic effects of aggregate demand and supply disturbances' *American Economic Review*, **79** (1989), 1146–64.

Box, G. E. P. and Jenkins, G. M., *Time-Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, 1970.

Cooley, T. F. and LeRoy, S. F., 'Atheoretical macroeconomics: a critique', *Journal of Monetary Economics*, **16** (1985) 283–308.

Curry, D. J., Divakar, S., Mathur, S.K. and Whiteman, C. H., 'BVAR as a category management tool: an illustration and comparison with alternative techniques', *Journal of Forecasting* (this issue).

Dickey, T. A. and Fuller, W. A., 'The likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica*, *49* (1981) 1057–72.

Doan, T., Litterman, R. and Sims, C., 'Forecasting and conditional projection using realistic prior distributions', *Econometric Reviews*, **3** (1984) 1–100.

Dua, P. and Ray, S. C., 'A BVAR Model for the Connecticut economy', *Journal of Forecasting* (this issue).

Dua, P. and Smyth, D. J., 'Forecasting US home sales using BVAR models and survey data on households buying attitudes for homes', *Journal of Forecasting* (this issue).

Engle, R. F. and Granger, C. W. J., 'Cointegration and error correction: representation, estimation and testing', *Econometrica*, **55** (1987) 251–76.

Engle, R. F. and Yoo, B. S., 'Forecasting and testing in cointegrated systems', *Journal of Econometrics*, **35** (1987), 143–59.
Gilbert, P. D., 'Combining VAR estimation and state space model reduction for simple good predictions', *Journal of Forecasting* (this issue).
Ioannidis, C., Laws, J., Matthews, K. and Morgan, B., 'Business cycle analysis and forecasting with a structural vector autoregression model for Wales', *Journal of Forecasting* (this issue).
Joutz, F. L., Maddala, G. S. and Trost, R. P., 'An integrated Bayesian vector autoregression and error correction model for forecasting electricity consumption and prices', *Journal of Forecasting* (this issue).
Kadiyala, K. R. and Karlsson, S., 'Forecasting with generalized Bayesian vector autoregressions', *Journal of Forecasting*, **12** (1993) 365–78.
Keating, J. W., 'Structural approaches to vector autoregressions', *Federal Reserve Bank of St Louis Review*, **74** (September/October 1992), 37–57.
Litterman, R. B., 'A statistical approach to economic forecasting', *Journal of Business and Economic Statistics*, **4** (1986) 1–4.
Lupoletti, W. M. and Webb, R. H., 'Defining and improving the accuracy of macroeconomic forecasts: contributions from a VAR model', *Journal of Business*, **59** (1986), 263–85.
McNees, S. K., 'Forecasting accuracy of alternative techniques: a comparison of US macroeconomic forecasts', *Journal of Business and Economic Statistics*, **4** (1986), 5–15.
Pinder, J. P. and Shoesmith, G. L., 'Modelling multivariate cointegrated systems: insights from nonlinear dynamics', *Journal of Forecasting* (this issue).
Robertson, D. and Wickens, M., 'VAR modelling' in Hall, S. (ed.), *Applied Economic Forecasting Techniques*, Hemel Hempstead: Harvester Wheatsheaf, 1994.
Sarantis, N. and Stewart, C., 'Structural, VAR and BVAR models of exchange rate determination: a comparison of their forecasting performance', *Journal of Forecasting* (this issue).
Shapiro, M. D. and Watson, M. W., 'Sources of business cycle fluctuations', in Fischer, S. (ed.), *NBER Macroeconomics Annual 1988*, Cambridge, MA: MIT Press, 1988.
Sims, C. A., 'Macroeconomics and reality' *Econometrica*, **48** (1980), 1–48.
Sims, C. A., 'Are forecasting methods usable for policy analysis?' *Federal Reserve Bank of Minneapolis Quarterly Review*, (Winter 1986), 2–16.
Sims, C. A., Stock, J. H. and Watson, M. W., 'Inference in linear time series models with some unit roots', *Econometrica* **55** (1990), 113–44.
Tiao, G .C. and Box, G. E. P., 'Modeling multiple time series with applications', *Journal of the American Statistical Association*, **76** (1981), 802–16.
Tinbergen, J., *Statistical Testing of Business Cycle Theories*, Vols I and II, Geneva: League of Nations, 1939.
Webb, R. H., 'Forecasts of inflation from VAR models', *Journal of Forecasting (this issue)*.
Whittle, P., *Prediction and Regulation by Linear Least Squares*, London: English Universities Press, 1963.
Wold, H., *A Study in the Analysis of Stationary Time Series*, Upsala: Almquist and Wiksell, 1938.
Zellner, A. and Palm, F., 'Time series analysis and simultaneous equation econometric models', *Journal of Econometrics*, **2** (1974), 17–54.

Author's biography:
Ken Holden is Professor of Applied Economics in the Business School at Liverpool John Moores University. He has a BSc from Leicester University, an MA(Econ) from Manchester University and is a Chartered Statistician. He has held appointments at the universities of Manchester, Liverpool and Otago and at the Open University. He is the author or joint author of seven books and over sixty papers on economic modelling and forecasting. He is an associate editor of the *Journal of Forecasting* and the *International Journal of Forecasting*.

Author's address:
Ken Holden, The Business School, Liverpool John Moores University, 98 Mount Pleasant, Liverpool L3 5UZ, U.K.