ASYMPTOTIC DISTRIBUTIONS OF IMPULSE RESPONSE FUNCTIONS AND FORECAST ERROR VARIANCE DECOMPOSITIONS OF VECTOR AUTOREGRESSIVE MODELS

Helmut Lütkepohl*

Abstract—In recent years vector autoregressive models have become standard tools for economic analyses. Impulse response functions and forecast error variance decompositions are usually computed from these models in order to investigate the interrelationships within the system. However, sometimes no measures of estimation uncertainty are provided by authors. One reason may be that the relevant asymptotic distribution theory is distributed over various publications. In this article the available results are summarized and the missing links are provided in order to facilitate the computation of standard errors and test statistics.

I. Introduction

OVER the last few years economic analyses based on vector autoregressive (VAR) models have become increasingly popular. Tracing the response of the system to an innovation in one of the variables and decomposing the forecast error variances have become standard tools for economic analyses.

In such analyses the data generation process is commonly assumed to be a K-dimensional VAR(p) process

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \tag{1}$$

where $y_t = (y_{1t}, \dots, y_{Kt})'$, the A_i are $(K \times K)$ coefficient matrices and $u_t = (u_{1t}, \dots, u_{Kt})'$ is K-dimensional white noise, that is, $E(u_t) = 0$ and

$$E(u_t u_s') = \begin{cases} \Sigma_u, & \text{if } t = s \\ 0 & \text{otherwise,} \end{cases}$$

and Σ_u is positive definite. In some analyses the VAR order p is allowed to be infinity.

The process (1) is assumed to be (covariance) stationary so that it has a moving average (MA) representation

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}, \tag{2}$$

where $\Phi_0 = I_K$ is the $(K \times K)$ identity matrix and

$$\Phi_i = \sum_{j=1}^{i} \Phi_{i-j} A_j, \; i = 1, 2, \dots, \tag{3}$$

with $A_j = 0$ for $j > p$. The elements of the Φ_i are sometimes interpreted as impulse responses of the system (e.g., Backus (1986)).

Many authors prefer to orthogonalize the u_t and then interpret the coefficients of the resulting MA representation (e.g., Sims (1980, 1981, 1986), Burbidge and Harrison (1984)). In other words, they choose a (usually triangular) matrix P with positive diagonal elements such that $\Sigma_u = PP'$ and define $w_t = P^{-1} u_t$, hence $\Sigma_w = E(w_t w_t') = I_K$. The corresponding MA representation of y_t is

$$y_t = \sum_{i=0}^{\infty} \Theta_i w_{t-i}, \tag{4}$$

where $\Theta_i = \Phi_i P$. For convenience P is assumed to be lower triangular in the following. Similar results also hold with other P matrices satisfying $PP' = \Sigma_u$.

Related quantities of interest are the accumulated responses

$$\Psi_j = \sum_{i=0}^{j} \Phi_i$$

and

$$\Xi_j = \sum_{i=0}^{j} \Theta_i = \Psi_j P.$$

Note that the total accumulated responses are

$$\Psi_\infty = \sum_{i=0}^{\infty} \Phi_i = \left(I_K - A_1 - \cdots - A_p\right)^{-1}$$

and $\Xi_\infty = \Psi_\infty P$. Furthermore, forecast error variance decompositions are often discussed in applied studies (e.g., Sims (1981), Runkle (1987)). The proportion of the h-step forecast error variance of variable k, accounted for by innovations

Received for publication June 9, 1988. Revision accepted for publication March 13, 1989.

*Christian-Albrechts-Universität Kiel.

The author thanks Knut Haase for carrying out the computations for the example in section IV.

in variable j, will be denoted by $\omega_{kj,h}$, that is,

$$\omega_{kj,h} = \sum_{i=0}^{h-1} \theta_{kj,i}^2 / MSE_k(h)$$
$$= \sum_{i=0}^{h-1} (e_k'\Theta_i e_j)^2 / MSE_k(h), \quad (5)$$

where $\theta_{kj,i}$ is the kj^{th} element of Θ_i, e_k is the k^{th} column of I_K and

$$MSE_k(h) = \sum_{i=0}^{h-1} e_k'\Phi_i\Sigma_u\Phi_i'e_k \quad (6)$$

is the mean squared error (forecast error variance) of an h-step forecast of variable k.

In practice, the A_i and Σ_u are unknown and must be estimated from the available data. In that case the impulse responses and forecast error variance components are usually computed from the estimated parameters. In contrast to common practice in other econometric work where t-ratios or standard errors are usually reported for estimated quantities, measures for estimation uncertainty are often *not* reported for impulse responses and forecast error variance components. The importance of such measures has been demonstrated by Runkle (1987) who shows that the estimation precision may be extremely low especially when unrestricted VAR models are fitted.

One reason for the silence on the estimation precision of the quantities of interest may be the fact that the relevant asymptotic distributions are not sufficiently easily accessible to applied researchers. Also statements about the difficulties in deriving general analytical results (see Baillie (1987, p. 111) or Runkle (1987, p. 438)) may be responsible for this state of affairs. The purpose of this article is to review and summarize the available results and to demonstrate that remarkably simple analytical expressions are available under assumptions often made in practice.

In fairness to previous authors it must be mentioned that in some articles standard errors are reported at least for some of the quantities of interest (e.g., Sims (1986), Burbidge and Harrison (1984), Runkle (1987)). Often Monte Carlo integration or bootstrap methods are used in those cases. These methods are computationally quite expensive. Moreover, the quality of the resulting estimators for the standard errors is not clear. Therefore we concentrate on more classical methods in the following, that is, we review the available asymptotic results. While the resulting estimators of the standard errors share the unknown small sample properties with those obtained by Monte Carlo or bootstrapping methods, the former are at least very easy to compute so that computational complexity can no longer be used as an excuse for not reporting them.

In the next section results for VAR processes with known orders are given and in section III some results for processes with unknown and possibly infinite orders are provided. An example is given in section IV, conclusions are drawn in section V, and the proof of the results in section II is provided in the appendix.

II. Results for VAR Processes with Known Order

We are mainly interested in providing simple expressions for the asymptotic standard errors of the impulse responses and forecast error variance components. For this purpose the following result from Serfling (1980, p. 122) is used. Suppose β is an $(n \times 1)$ vector of parameters and $\hat{\beta}$ is an estimator such that

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma_\beta), \quad (7)$$

where $\xrightarrow{d}$ denotes convergence in distribution, $N(0, \Sigma_\beta)$ denotes the multivariate normal distribution with mean vector 0 and covariance matrix Σ_β and T is the sample size used for estimation. Furthermore, let $g(\beta) = (g_1(\beta), \ldots, g_m(\beta))'$ be a continuously differentiable function with values in m-dimensional Euclidean space and $\partial g_i/\partial\beta' = (\partial g_i/\partial\beta_j)$ is nonzero at β for $i = 1, \ldots, m$. Then

$$\sqrt{T}\left[g(\hat{\beta}) - g(\beta)\right] \xrightarrow{d} N\left(0, \frac{\partial g}{\partial\beta'}\Sigma_\beta\frac{\partial g'}{\partial\beta}\right). \quad (8)$$

This result has been used in the literature in order to establish the asymptotic normal distribution of the dynamic multipliers. Runkle (1987) suggests using numerical methods to compute the derivatives $\partial g/\partial\beta'$. However, it is not difficult to derive analytical expressions for these derivatives. They will be given in Proposition 1.

In writing down the results formally we use the following notation in addition to that defined in section I:

$$\alpha = \operatorname{vec}(A_1, \ldots, A_p) \qquad (K^2 p \times 1)$$

$$A = \begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & I_K & 0 \end{bmatrix} (Kp \times Kp)$$

$$\sigma = \operatorname{vech}(\Sigma_u) \qquad (K(K+1)/2 \times 1)$$

and the corresponding estimators are furnished with a caret. Here vec denotes the column stacking operator and vech is the corresponding operator that stacks only the elements on and below the diagonal. This operator corresponds to choosing a lower triangular matrix P in section I such that $PP' = \Sigma_u$. If an upper triangular matrix P is chosen, the vech operator or some of the following results have to be modified.

As usual the Kronecker product is denoted by $\otimes$, the commutation matrix K_{mn} is defined such that, for any $(m \times n)$ matrix G, $K_{mn} \operatorname{vec}(G) = \operatorname{vec}(G')$, and the $(m^2 \times m(m+1)/2)$ duplication matrix D_m is defined such that $D_m \operatorname{vech}(F) = \operatorname{vec}(F)$ for a symmetric $(m \times m)$ matrix F (e.g., Magnus and Neudecker (1986)). Furthermore L_m is the $(m(m+1)/2 \times m^2)$ elimination matrix defined such that, for any $(m \times m)$ matrix F, $\operatorname{vech}(F) = L_m \operatorname{vec}(F)$, and $J = [I_K \; 0 \ldots 0]$ is a $(K \times Kp)$ matrix. With this notation Proposition 1 can be stated.

PROPOSITION 1:

Suppose

$$\sqrt{T}\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\sigma} - \sigma \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma_\sigma \end{bmatrix}\right).$$

Then

$$\sqrt{T}\operatorname{vec}(\hat{\Phi}_i - \Phi_i) \xrightarrow{d} N(0, G_i \Sigma_\alpha G_i'), \qquad i = 1, 2, \ldots, \qquad \text{(i)}$$

where

$$G_i = \partial \operatorname{vec}(\Phi_i)/\partial \alpha' = \sum_{m=0}^{i-1} J(A')^{i-1-m} \otimes \Phi_m;$$

$$\sqrt{T}\operatorname{vec}(\hat{\Psi}_j - \Psi_j) \xrightarrow{d} N(0, F_j \Sigma_\alpha F_j'), \qquad j = 1, 2, \ldots, \qquad \text{(ii)}$$

where $F_j = G_1 + \cdots + G_j$;

$$\sqrt{T}\operatorname{vec}(\hat{\Psi}_\infty - \Psi_\infty) \xrightarrow{d} N(0, F_\infty \Sigma_\alpha F_\infty'), \qquad \text{(iii)}$$

where $F_\infty = \underbrace{(\Psi_\infty', \ldots, \Psi_\infty')}_{p \text{ times}} \otimes \Psi_\infty$;

$$\sqrt{T}\operatorname{vec}(\hat{\Theta}_i - \Theta_i) \xrightarrow{d} N(0, C_i \Sigma_\alpha C_i' + \bar{C}_i \Sigma_\sigma \bar{C}_i'), \qquad i = 0, 1, 2, \ldots, \qquad \text{(iv)}$$

where

$C_0 = 0, \quad C_i = (P' \otimes I_K) G_i, \quad i = 1, 2, \ldots,$
$\bar{C}_i = (I_K \otimes \Phi_i) H, i = 0, 1, \ldots,$ and

$$H = \partial \operatorname{vec}(P)/\partial \sigma' = L_K' \{ L_K (I_{K^2} + K_{KK})(P \otimes I_K) L_K' \}^{-1};$$

$$\sqrt{T}\operatorname{vec}(\hat{\Xi}_j - \Xi_j) \xrightarrow{d} N(0, B_j \Sigma_\alpha B_j' + \bar{B}_j \Sigma_\sigma \bar{B}_j'), \qquad j = 1, 2 \ldots \qquad \text{(v)}$$

where $B_j = (P' \otimes I_K) F_j$ and $\bar{B}_j = (I_K \otimes \Psi_j) H$;

$$\sqrt{T}\operatorname{vec}(\hat{\Xi}_\infty - \Xi_\infty) \xrightarrow{d} N(0, B_\infty \Sigma_\alpha B_\infty' + \bar{B}_\infty \Sigma_\sigma \bar{B}_\infty'), \qquad \text{(vi)}$$

where

$B_\infty = (P' \otimes I_K) F_\infty$ and $\bar{B}_\infty = (I_K \otimes \Psi_\infty) H$;

$$\sqrt{T}(\hat{\omega}_{kj,h} - \omega_{kj,h}) \xrightarrow{d} N(0, d_{kj,h} \Sigma_\alpha d_{kj,h}' + \bar{d}_{kj,h} \Sigma_\sigma \bar{d}_{kj,h}') \qquad k, j = 1, \ldots, K; \quad h = 1, 2, \ldots, \qquad \text{(vii)}$$

where $d_{kj,1} = 0$ and for $h > 1$.

$$d_{kj,h} = 2 \sum_{i=1}^{h-1} \left[MSE_k(h)(e_k' \Phi_i P e_j)(e_j' P' \otimes e_k') G_i - (e_k' \Phi_i P e_j)^2 \sum_{m=1}^{h-1} (e_k' \Phi_m \Sigma_u \otimes e_k') G_m \right] / MSE_k(h)^2$$

and, for $h = 1, 2, \ldots,$

$$\bar{d}_{kj,h} = \sum_{i=0}^{h-1} \left[2\, MSE_k(h)(e_k' \Phi_i P e_j)(e_j' \otimes e_k' \Phi_i) H - (e_k' \Phi_i P e_j)^2 \sum_{m=0}^{h-1} (e_k' \Phi_m \otimes e_k' \Phi_m) D_k \right] / MSE_k(h)^2.$$

In the appendix references for proofs and the missing links are given. From this proposition approximate variances of the estimated impulse responses and forecast error variance components are obtained in the usual way by dividing the diagonal elements of the asymptotic covariance matrices by the sample size T. For instance, denoting the kj^{th} element of $\hat{\Phi}_i$ by $\hat{\varphi}_{kj,i}$, the approximate variance of $\hat{\varphi}_{11,i}$ is the upper left-hand corner element of $G_i\Sigma_\alpha G_i'/T$. Thus, as usual the approximate variances go to zero with the sample size. Some remarks regarding the proposition are now worthwhile.

(i) In the proposition some matrices of partial derivatives may be zero. For instance, if a VAR(1) model is fitted although the true order is zero, that is, y_t is white noise, then $G_2 = 0$ and a degenerate asymptotic distribution with zero covariance matrix is obtained for $\sqrt{T}(\hat{\Phi}_2 - \Phi_2)$. For simplicity we call such a distribution also multivariate normal. Otherwise it would be necessary to distinguish between cases with zero and nonzero asymptotic variances. Notice, however, that estimating the asymptotic variances in the usual way by replacing the unknown quantities by estimators may be inappropriate in this case. As a consequence a confidence interval of an impulse response coefficient with a zero asymptotic variance may not have the desired asymptotic probability content.

(ii) If the VAR(p) process y_t is (covariance) stationary with

$$\det\left(I_K - A_1 z - \cdots - A_p z^p\right) \neq 0 \text{ for } |z| \leq 1,$$

and the u_t are independently, identically distributed (iid) with bounded fourth moments, then the usual LS estimators have asymptotic covariance matrix

$$\Sigma_\alpha = \Gamma^{-1} \otimes \Sigma_u, \tag{9}$$

where

$$\Gamma = E\left(\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} [y_t', \ldots, y_{t-p+1}']\right).$$

If y_t is Gaussian, $\hat{\alpha}$ and $\hat{\sigma}$ are asymptotically independent as assumed in Proposition 1 with Σ_α as in (9) and, defining $D_K^+ = (D_K' D_K)^{-1} D_K'$,

$$\Sigma_\sigma = 2\, D_K^+ (\Sigma_u \otimes \Sigma_u) D_K^{+\prime}.$$

However, the proposition can also be used if subset VAR models (e.g., Penm and Terrell (1982, 1984)) or VAR models with nonlinear parameter restrictions (e.g., Reinsel (1983), Velu, Reinsel and Wichern (1986)) are fitted. In fact, even if vector autoregressive moving average (VARMA) models are fitted the proposition will be useful, if the asymptotic distribution of the corresponding pure VAR representation can be found.

(iii) If the parameters of an unrestricted VAR model are estimated equation by equation it may be easier to think in terms of the parameter vector

$$a = \operatorname{vec}\left[(A_1, \ldots, A_p)'\right].$$

Denoting the covariance matrix of the asymptotic distribution of the corresponding estimator $\hat{a}$ by Σ_a, we get $\Sigma_\alpha = Q\Sigma_a Q'$, where $Q = \partial\alpha/\partial a'$ is a matrix consisting of unit vectors in such a way that $\alpha = Qa$.

(iv) In (1) y_t is implicitly assumed to have zero mean. This assumption was made for convenience. Proposition 1 remains unaltered if a nonzero mean term, a polynomial trend or a seasonal component is removed prior to estimating the VAR parameters. Alternatively, polynomial trend terms and seasonal dummies may be included in the model (1) and estimated jointly with the VAR coefficients without affecting Proposition 1. This result follows obviously from the fact that the quantities of interest in Proposition 1 only depend on the A_i and Σ_u.

(v) Test statistics for testing hypotheses that involve several of the response coefficients or forecast error variance components can be obtained from Proposition 1 in the usual way. However, it has to be taken into account that, for instance, the elements of $\hat{\Phi}_i$ and $\hat{\Phi}_j$ will not be independent asymptotically. If elements from two or more matrices are involved in the null hypothesis the joint distribution of all the matrices must be determined. This distribution can be derived easily from the results given in Proposition 1. For instance, the covariance matrix of the joint asymptotic distribution of $\operatorname{vec}(\hat{\Phi}_i, \hat{\Phi}_j)$ is

$$\frac{\partial \operatorname{vec}(\Phi_i, \Phi_j)}{\partial \alpha'} \Sigma_\alpha \frac{\partial \operatorname{vec}(\Phi_i, \Phi_j)'}{\partial \alpha},$$

where

$$\partial \operatorname{vec}(\Phi_i, \Phi_j)/\partial \alpha' = \begin{bmatrix} \partial \operatorname{vec}(\Phi_i)/\partial \alpha' \\ \partial \operatorname{vec}(\Phi_j)/\partial \alpha' \end{bmatrix},$$

etc. We have chosen to state the proposition for individual MA coefficient matrices because that way all required matrices have relatively small dimensions and hence are easy to compute.

(vi) Summing the forecast error variance components over j,

$$\sum_{j=1}^{K} \omega_{kj,h} = \sum_{j=1}^{K} \hat{\omega}_{kj,h} = 1$$

for each k and h. These restrictions are not taken into account in the derivation of the asymptotic distributions of these components in (vii). For $K = 1$, the standard errors obtained from Proposition 1 are zero, however, as they should be. Note also that the asymptotic variance of $\hat{\omega}_{kj,h}$ is zero if $\omega_{kj,h} = 0$ since in this case $d_{kj,h}$ and $\bar{d}_{kj,h}$ are both zero. Thus, in this case, the asymptotic distribution cannot be used in the usual way for tests of significance and for setting up confidence intervals. This state of affairs is unfortunate because testing the significance of forecast error variance components is of particular interest in practice.

(vii) In section I it was mentioned that y_t is assumed to be stationary. Otherwise the MA representation will in general not exist. However, if the deviations from some equilibrium relationship are stationary like, for example, in cointegrated systems (Engle and Granger (1987)), Proposition 1 may be used to obtain the asymptotic distributions of the responses of the system to disturbances of the equilibrium.

In practice, the order of the underlying VAR process will usually be unknown. Some consequences for the impulse responses will be summarized in the following section.

III. Unknown VAR Order

If the data are generated by a stationary VAR process of unknown order, asymptotic properties of the impulse responses of the system can be derived under the following two alternative assumptions:

(i) An upper bound for the VAR order is known.
(ii) The order of the VAR process that is fitted to the given multiple time series approaches infinity with the time series length.

If the first assumption is made one could, of course, simply use the upper bound as the true VAR order because zero coefficient matrices are not prohibited in Proposition 1. In other words, if a coefficient matrix is not known to be zero, it is estimated along with the other parameters. Although this approach is theoretically sound, it is not quite satisfactory from a practical point of view because the estimation uncertainty grows with the number of parameters. Also, as mentioned previously, zero asymptotic variances may result that cannot be used in the usual way to set up confidence intervals or test statistics. Therefore, it is desirable to estimate as few VAR coefficients as possible. In other words, it is desirable to find the smallest upper bound for the VAR order. This may be done with the help of VAR order selection criteria such as AIC, HQ or SC (e.g., Lütkepohl (1985)). If a consistent criterion like HQ or SC is used, the asymptotic distributions of the impulse responses and forecast error variance components are the same as in the case of a known VAR order.

If the orders of the processes fitted to the data are assumed to approach infinity with the sample size, the approach of Lütkepohl (1988) may be used to obtain some asymptotic results. In this approach y_t is permitted to have infinite VAR order, that is,

$$y_t = \sum_{i=1}^{\infty} A_i y_{t-i} + u_t. \tag{10}$$

To ensure that this notation is meaningful, the A_i are assumed to be absolutely summable so that

$$\sum_{i=1}^{\infty} \|A_i\| < \infty, \tag{11}$$

where $\|A_i\|^2 = \mathrm{tr}(A_i'A_i)$. Of course, the A_i may be zero for i greater than some finite number p. If y_t is a stationary and invertible VARMA process, condition (11) is satisfied.

It is assumed that, although the VAR order of y_t is potentially infinite, only finite order VAR(p) models are fitted to a finite multiple time series of length T using multivariate LS estimation. The relation between p and T is specified in the following assumption.

ASSUMPTION 1:

The order p of a VAR model fitted to a multiple time series of length T depends on T such that $p \to \infty$, $p^3/T \to 0$, and

$$\sqrt{T}\left(\sum_{i=p+1}^{\infty} \|A_i\|\right) \to 0 \tag{12}$$

as $T \to \infty$.

This condition requires that the VAR order goes to infinity at a much slower rate than the sample size because $p^3/T \to 0$ which imposes an upper bound for the rate at which p is permitted to go to infinity. In contrast (12) is a lower bound for that rate. A further discussion of Assumption 1 may be found in Lütkepohl (1987, sec. 2.4.3).

The MA coefficient matrices Φ_i of the representation (2) may be computed recursively using (3). The corresponding MA coefficient estimators based on an estimated VAR(p) process are

$$\tilde{\Phi}_{i,p} = \sum_{j=0}^{i-1} \tilde{\Phi}_{j,p}\tilde{A}_{i-j,p}, \qquad i = 1,2,\ldots, \tag{13}$$

where $\tilde{A}_{j,p}$ is the multivariate LS estimator for $j \leq p$ and $\tilde{A}_{j,p} = 0$ for $j > p$. Furthermore we define $\tilde{\Psi}_{j,p} = I_K + \tilde{\Phi}_{1,p} + \cdots + \tilde{\Phi}_{j,p}$. The asymptotic distributions of $\tilde{\Phi}_{i,p}$ and $\tilde{\Psi}_{j,p}$ are given in the next proposition.

PROPOSITION 2:

Let y_t be a stationary process with VAR representation (10), where $u_t = (u_{1t}, \ldots, u_{Kt})'$ is iid white noise with positive definite covariance matrix $\Sigma_u = E(u_t u_t')$ and

$$E|u_{it}u_{jt}u_{mt}u_{nt}| < \infty \quad \text{for } 1 \leq i, j, m, n \leq K.$$

Furthermore, suppose that (11) holds, $\det(I_K - \Sigma_{i=1}^{\infty} A_i z^i) \neq 0$ for $|z| \leq 1$, and Assumption 1 is satisfied. Then,

$$\sqrt{T}\,\mathrm{vec}\left(\tilde{\Phi}_{i,p} - \Phi_i\right) \xrightarrow{d} N(0, \Omega_i),$$

where

$$\Omega_i = \Sigma_u^{-1} \otimes \sum_{m=0}^{i-1} \Phi_m \Sigma_u \Phi_m'$$

and

$$\sqrt{T}\,\mathrm{vec}\left(\tilde{\Psi}_{j,p} - \Psi_j\right) \xrightarrow{d} N\left(0, \bar{\Omega}_j\right),$$

where

$$\bar{\Omega}_j = \Sigma_u^{-1} \otimes \sum_{n=1}^{j} \sum_{m=1}^{j} \sum_{i=0}^{m-1} \Phi_i \Sigma_u \Phi_{n-m+i}'$$

with $\Phi_i = 0$ for $i < 0$.

A proof of Proposition 2 is given by Lütkepohl (1988). The covariance matrices of the asymptotic distributions now look even simpler than in Proposition 1(i), (ii). However, if the true VAR order is less than p and a full VAR(p) model is fitted to the data so that $\Sigma_\alpha = \Gamma^{-1} \otimes \Sigma_u$ as in (9), then Proposition 1(i) gives

$$\begin{aligned} G_1 \Sigma_\alpha G_1' &= (J \otimes \Phi_0)\left(\Gamma^{-1} \otimes \Sigma_u\right)(J \otimes \Phi_0)' \\ &= J\Gamma^{-1}J' \otimes \Sigma_u = \Sigma_u^{-1} \otimes \Sigma_u = \Omega_1 \end{aligned}$$

(see Lütkepohl (1988)). Thus, the two asymptotic covariance matrices for the estimator of Φ_1 that are given in Propositions 1 and 2 are identical. On the other hand, different covariance matrices are obtained for the estimators of Φ_i, $i > 1$. Obviously the diagonal elements of Ω_i are nondecreasing with growing i while the variances of the elements of Φ_i may decrease with increasing i under the conditions of Proposition 1. The reason is that, under the conditions of Proposition 2, the order of the VAR models fitted to the data approaches infinity even if the actual order is finite. Thus, more and more parameters are estimated as T goes to infinity.

The asymptotic covariance matrices of the Θ_i, Ξ_j and the forecast error variance components seem to be unknown under the assumptions of Proposition 2. In the next section we illustrate the foregoing results by an example.

IV. Example

As an example we consider a three-dimensional system consisting of first differences of logarithms of quarterly, seasonally adjusted West German fixed investment (y_1), disposable income (y_2) and consumption expenditures (y_3), all in current prices. We use the data from 1960.I to 1978.IV as published by the Deutsche Bundesbank. In an investigation of structural changes a possible structural break was found in 1979 (Lütkepohl (1989a)). Since nonstationarity of this type may invalidate the asymptotic distributions in Propositions 1 and 2 we decided to use data up to 1978 only. First differences of logarithms (rates of change) were used to remove trends. Such a transformation would be problematic if the variables were cointegrated. However, since Proposition 2 requires stationarity and the example is meant to illustrate the two asymptotic theories of the previous sections we have chosen to use rates of change rather than levels of the variables.

Allowing for a maximum VAR order of eight, Akaike's AIC criterion was minimized for order two while HQ and SC both chose order zero. We use order $p = 2$ in the following because SC and to some extent also HQ have some probability of underestimating the true VAR order in small sam-

FIGURE 1.—RESPONSES OF CONSUMPTION TO A UNIT SHOCK IN INCOME WITH TWO-STANDARD ERROR BOUNDS

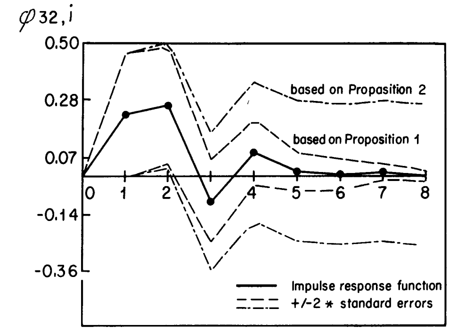


FIGURE 2.—ACCUMULATED RESPONSES OF CONSUMPTION TO UNIT SHOCK IN INCOME WITH TWO-STANDARD ERROR BOUNDS

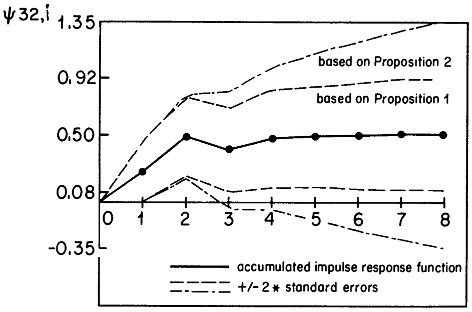


FIGURE 3.—RESPONSES OF CONSUMPTION TO INCOME INNOVATION WITH TWO-STANDARD ERROR BOUNDS BASED ON PROPOSITION 1

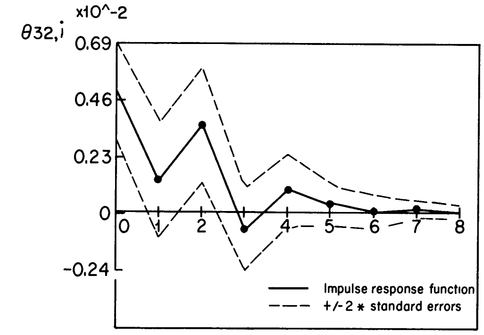
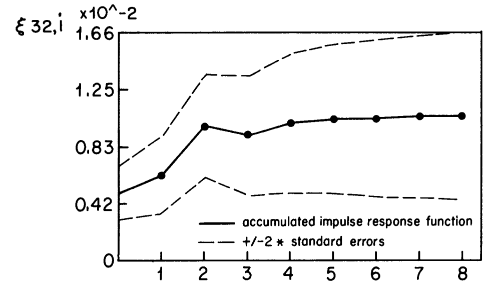


FIGURE 4.—ACCUMULATED RESPONSES OF CONSUMPTION TO INCOME INNOVATION WITH TWO-STANDARD ERROR BOUNDS BASED ON PROPOSITION 1



ples when the dimension of the system is moderate or large. Note that underestimation of the order is not permitted in Proposition 1. Of course, a VAR(0) process would be useless for illustrative purposes because its impulse responses are all zero. If in practice the order is chosen too large this may result in imprecise coefficient estimates and may be reflected in large standard errors of the impulse responses. The VAR(2) model was estimated by multivariate least squares. Intercept terms were estimated jointly with the VAR parameters rather than using mean corrected data.

In figures 1–4 selected impulse responses with two-standard error bounds are shown. In figures 1 and 2 bounds from Propositions 1 and 2 are given. The computations were performed using the computer language GAUSS which is especially suitable because it supports matrix computations. Obviously the estimation uncertainty is quite substantial. The standard errors obtained from Proposition 2 are much larger than those based on Proposition 1 for longer lead times. This reflects the fact that the VAR order is assumed to go to infinity with the sample size in Proposition 2 so that, even for longer lead times, the estimation uncertainty does not vanish.

On the other hand, the standard errors of the $\hat{\Phi}_i$ and $\hat{\Theta}_i$ impulses based on Proposition 1 must go to zero as $i \to \infty$ because, for stationary processes, the Φ_i and Θ_i (and A^i) go to zero as $i \to \infty$ and these terms are essential ingredients of the asymptotic covariance matrices in Proposition 1(i), (iv). Intuitively, if the data generation process is known to be a stationary, finite order VAR(p) process the impulse responses are known to taper off to zero after some periods and hence have a large probability of being close to zero for large i. This is reflected in the confidence bounds becom-

TABLE 1.—FORECAST ERROR VARIANCE DECOMPOSITION OF THE INVESTMENT / INCOME / CONSUMPTION SYSTEM WITH STANDARD ERRORS IN PARENTHESES

Forecast Error in	Forecast Horizon h	Proportions of Forecast Error Variance h Periods Ahead Accounted for by Innovations in		
		Investment	Income	Consumption
Investment	1	1.00(.00)	.00(.00)	.00(.00)
	2	.96(.04)	.02(.03)	.02(.03)
	3	.95(.04)	.03(.03)	.03(.03)
	4	.94(.05)	.03(.03)	.03(.03)
	8	.94(.05)	.03(.03)	.03(.04)
Income	1	.02(.04)	.98(.04)	.00(.00)
	2	.06(.06)	.91(.07)	.03(.04)
	3	.07(.06)	.90(.07)	.03(.04)
	4	.07(.06)	.89(.07)	.04(.04)
	8	.07(.06)	.89(.07)	.04(.04)
Consumption	1	.08(.09)	.27(.14)	.65(.14)
	2	.08(.08)	.27(.13)	.65(.13)
	3	.13(.10)	.33(.13)	.54(.13)
	4	.13(.10)	.34(.13)	.54(.12)
	8	.13(.10)	.34(.13)	.53(.12)

ing smaller. In contrast, the two-standard error bounds in figures 2 and 4 do not approach zero because the $\Psi_j = (\psi_{kh,j})$ and $\Xi_j = (\xi_{kh,j})$ are accumulated impulse responses that do not approach zero for $j \to \infty$ in general. Intuitively speaking, adding another random variable to a sum of random variables here leads to an increase in the variance.

Although the estimation uncertainty is substantial, it is worth noting that qualitatively the impulse responses are as expected, that is, they have the expected sign. Large estimation uncertainty is the common price that has to be paid in VAR analyses for not forcing possibly false a priori structure on the system.

In table 1 forecast error variance decompositions are given. The standard errors obtained from Proposition 1 are quite large so that the forecast error variance of investment accounted for by income or consumption may not be significantly different from zero for any of the forecast horizons shown. One interpretation of this result is that investment may be exogenous in this system. Of course, the system is not meant to explain the structure of investment in the West German economy. Thus the result is not surprising. It is also expected that income innovations account for some of the consumption variability. Under a two standard error criterion these are, in fact, the only significant contributions of one variable to the forecast error variance of another variable in table 1. Unfortunately, as mentioned in section II(vi), a formal significance test for a forecast error variance component is not possible on the basis of Proposition 1(vii).

V. Conclusions

In this paper the asymptotic distributions of dynamic responses and forecast error variance components computed from estimated VAR models are surveyed. Results are summarized for two scenarios. First it is assumed that the order of the data generation process is known and finite. The second scenario permits the VAR order to be unknown and even allows the true order to be infinite as may be the case when the data generation process is actually a VARMA process. An example is used to demonstrate both the applicability of the results in practice and the importance of attaching measures of estimation uncertainty to the estimates of impulse responses and forecast error variance components. In particular, the example shows that the usual estimates based on a full VAR(p) model may be quite unreliable.

In practice, Monte Carlo or bootstrap methods are sometimes used to assess the sampling variability of the quantities of interest in this survey. Currently it is not clear whether these methods provide more reliable standard errors than the asymptotic theory discussed in the foregoing. On the other hand, the standard errors based on asymptotic theory are much more efficient computationally than those obtained with resampling methods.

APPENDIX

Proof of Proposition 1

Baillie (1987) proves (i), and (ii) is an immediate consequence because

$$F_j = \partial \operatorname{vec}(I_K + \Phi_1 + \cdots + \Phi_j)/\partial\alpha' = \sum_{i=1}^{j} \partial \operatorname{vec}(\Phi_i)/\partial\alpha'.$$

Part (iii) follows by noting that

$$F_\infty = \frac{\partial \operatorname{vec}(\Psi_\infty)}{\partial\alpha'} = \frac{\partial \operatorname{vec}(\Psi_\infty)}{\partial \operatorname{vec}(\Psi_\infty^{-1})'} \frac{\partial \operatorname{vec}(\Psi_\infty^{-1})}{\partial\alpha'} = -(\Psi_\infty' \otimes \Psi_\infty)\frac{\partial \operatorname{vec}(I_K - A_1 - \cdots - A_p)}{\partial\alpha'}$$

(see Magnus and Neudecker (1986, equation (88))). Part (iv) is shown by Lütkepohl (1989b) and (v) is obtained by noting that

$$B_j = \partial \operatorname{vec}(\Psi_j P)/\partial\alpha' = (P' \otimes I_K)\, \partial \operatorname{vec}(\Psi_j)/\partial\alpha',$$

$$\bar{B}_j = \partial \operatorname{vec}(\Psi_j P)/\partial\sigma' = (I_K \otimes \Psi_j)\, \partial \operatorname{vec}(P)/\partial\sigma',$$

and $\partial \operatorname{vec}(P)/\partial\sigma' = H$ as shown in Lütkepohl (1989b). The result in (vi) follows in a similar way.

Finally (vii) is obtained since

$$\partial\omega_{kj,h}/\partial\alpha' = \left(2\sum_{i=0}^{h-1} (e_k'\Phi_i P e_j)(e_j'P' \otimes e_k') \times \frac{\partial \operatorname{vec}(\Phi_i)}{\partial\alpha'} MSE_k(h) - \sum_{i=0}^{h-1} (e_k'\Phi_i P e_j)^2 \frac{\partial MSE_k(h)}{\partial\alpha'}\right) \Bigg/ [MSE_k(h)]^2$$

$$\begin{aligned}\partial MSE_k(h)/\partial\alpha' &= \sum_{m=0}^{h-1} \big((e_k'\Phi_m\Sigma_u \otimes e_k')\,\partial \operatorname{vec}(\Phi_m)/\partial\alpha' + (e_k' \otimes e_k'\Phi_m\Sigma_u)\,\partial \operatorname{vec}(\Phi_m')/\partial\alpha'\big)\\ &= \sum_{m=1}^{h-1} [(e_k'\Phi_m\Sigma_u \otimes e_k') + (e_k' \otimes e_k'\Phi_m\Sigma_u)K_{KK}] \times \partial \operatorname{vec}(\Phi_m)/\partial\alpha'\\ &= \sum_{m=1}^{h-1} [(e_k'\Phi_m\Sigma_u \otimes e_k') + K_{11}(e_k'\Phi_m\Sigma_u \otimes e_k')]G_m\\ &\quad \text{(see Magnus and Neudecker (1986, Lemma 4))}\\ &= 2\sum_{m=1}^{h-1} (e_k'\Phi_m\Sigma_u \otimes e_k')G_m,\end{aligned}$$

$$\partial\omega_{kj,h}/\partial\sigma' = \sum_{i=0}^{h-1} \left(2(e_k'\Phi_i P e_j)(e_j' \otimes e_k'\Phi_i)\frac{\partial \operatorname{vec}(P)}{\partial\sigma'} MSE_k(h) - (e_k'\Phi_i P e_j)^2 \frac{\partial MSE_k(h)}{\partial\sigma'}\right) \Bigg/ [MSE_k(h)]^2$$

and

$$\begin{aligned}\partial MSE_k(h)/\partial\sigma' &= \sum_{m=0}^{h-1} (e_k'\Phi_m \otimes e_k'\Phi_m)\,\partial \operatorname{vec}(\Sigma_u)/\partial\sigma'\\ &= \sum_{m=0}^{h-1} (e_k'\Phi_m \otimes e_k'\Phi_m)\, D_K\, \partial \operatorname{vech}(\Sigma_u)/\partial\sigma'.\end{aligned}$$

Thereby Proposition 1 is proven.

REFERENCES

Backus, David, "The Canadian–U.S. Exchange Rate: Evidence from a Vector Autoregression," this REVIEW 68 (Nov. 1986), 628–637.

Baillie, Richard T., "Inference in Dynamic Models Containing 'Surprise' Variables," *Journal of Econometrics* 35 (1987), 101–117.

Burbidge, John, and Alan Harrison, "Testing for the Effect of Oil-Price Rises Using Vector Autoregressions," *International Economic Review* 25 (June 1984), 459–484.

Engle, Robert F., and Clive W. J. Granger, "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55 (Mar. 1987), 251–276.

Lütkepohl, Helmut, "Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process," *Journal of Time Series Analysis* 6 (1985), 35–52, "Correction," 8 (1987), 373.

______, *Forecasting Aggregated Vector ARMA Processes* (Berlin: Springer Verlag, 1987).

______, "Asymptotic Distribution of the Moving Average Coefficients of an Estimated Vector Autoregressive Process," *Econometric Theory* 4 (Apr. 1988), 77–85.

______, "Prediction Tests for Structural Stability of Multiple Time Series," *Journal of Business & Economic Statistics* 7 (Jan. 1989a), 129–135.

______, "A Note on the Asymptotic Distribution of Impulse Response Functions of Estimated VAR Models with Orthogonal Residuals," *Journal of Econometrics* (1989b), forthcoming.

Magnus, Jan R., and Heinz Neudecker, "Symmetry, 0–1 Matrices and Jacobians: A Review," *Econometric Theory* 2 (Aug. 1986), 157–190.

Penm, Jack H. W., and R. D. Terrell, "On the Recursive Fitting of Subset Autoregressions," *Journal of Time Series Analysis* 3 (1982), 43–59.

———, "Multivariate Subset Autoregressive Modelling with Zero Constraints for Detecting 'Overall Causality'," *Journal of Econometrics* 24 (1984), 311–330.

Reinsel, Gregory, "Some Results on Multivariate Autoregressive Index Models," *Biometrika* 70 (Apr. 1983), 145–156.

Runkle, David E., "Vector Autoregressions and Reality," *Journal of Business & Economic Statistics* 5 (Oct. 1987), 437–442.

Serfling, Robert J., *Approximation Theorems of Mathematical Statistics* (New York: John Wiley, 1980).

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica* 48 (Jan. 1980), 1–48.

———, "An Autoregressive Index Model for the U.S. 1948–1975," in Jan Kmenta and J. B. Ramsey (eds.), *Large-Scale Macro-Econometric Models* (Amsterdam: North-Holland, 1981), 283–327.

———, "Are Forecasting Models Usable for Policy Analysis?" Federal Reserve Bank of Minneapolis, *Quarterly Review* (Winter 1986), 2–16.

Velu, Raja P., Gregory C. Reinsel, and Dean W. Wichern, "Reduced Rank Models for Multiple Time Series," *Biometrika* 73 (Apr. 1986), 105–118.