Journal of Econometrics 59 (1993) 319–341. North-Holland

Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions*

Phillip A. Braun
Northwestern University, Evanston, IL 60201, USA

Stefan Mittnik
State University of New York, Stony Brook, NY 11794, USA

Received July 1989, final version received June 1992

This paper investigates the consequences of model misspecifications in vector autoregressions (VARs). We specifically consider the effects of misspecifications on estimated impulse responses and variance decompositions. Assuming the underlying economic system has an ARMA representation, we derive closed-form expressions for the inconsistency in VAR-parameter estimates caused by omitted variables, ignored moving average terms, incorrectly specified lag lengths, or incorrect orthogonalization of innovations. Using these results we derive the inconsistencies for implied impulse responses and variance decompositions. Employing a trivariate economic model, we examine the potential seriousness of these misspecifications in applied work.

1. Introduction

In this paper we discuss the implications of fitting a vector autoregression (VAR) to variables (or a subset of variables) generated by a vector autoregressive moving average (ARMA) process and demonstrate the effects of potential misspecification due to the restrictions that are inherent in the VAR approach. Specific sources of misspecifications are considered. First, given the degrees-of-freedom constraints, VAR models usually contain only a limited number of variables, even though the analyst expects the underlying process to include additional variables. Second, the autoregressive order of the underlying process

Correspondence to: Stefan Mittnik, Department of Economics, State University of New York, Stony Brook, NY 11794-4384, USA.

* The authors would like to thank J. Caskey, E. Greenberg, L.H. Meyer, P. Rossi, two anonymous referees, and, especially, A. Zellner and the associate editor for valuable comments on earlier versions of this paper. Remaining errors are the authors' responsibility.

is underestimated. Third, if the disturbances of the underlying process are serially correlated, fitting a low-order or any finite-order VAR model may be inappropriate. These three types of misspecifications will generally affect the estimated VAR parameters and, therefore, affect derived impulse responses and variance decompositions. In addition, we theoretically investigate the consequences of a fourth type of misspecification. To use VARs for policy analysis, Sims (1980) orthogonalized the VAR residuals by decomposing the estimated residual-covariance matrix using the Cholesky decomposition.[1] We examine the effects of incorrect orthogonalizations on impulse responses and variance decompositions.

The paper is organized as follows. In section 2 we discuss implications of approximating a vector ARMA process, or a subset of variables generated by a vector ARMA process, with a VAR. In section 3 we derive closed-form expressions for the inconsistency of estimated impulse responses and variance decompositions due to VAR misspecifications. In section 4, using a trivariate macroeconomic system, we present quantitative examples to demonstrate the extent to which some of these misspecifications can distort policy analyses with VARs.

2. Fitting misspecified vector autoregressions

In this section we derive the inconsistency in VAR-parameter estimates that arise when the underlying system follows a vector ARMA(p, q) process, but the analyst fits a VAR of order k. By inconsistency we mean that the estimator does not converge to the true value of the corresponding coefficient of the process specified by representation (1) below.[2]

We assume the economic process of interest can be represented by the time-invariant, zero mean, weakly stationary, invertible, h-dimensional ARMA process

$$A(L)z_t = B(L)\varepsilon_t, \quad \mathrm{E}(\varepsilon_s\varepsilon_t') = \delta_{st}\Sigma, \quad \varepsilon_t = \Sigma^{1/2}v_t, \tag{1}$$

where $A(L) = I - A_1L - \cdots - A_pL^p$ and $B(L) = I + B_1L + \cdots + B_qL^q$ are $h \times h$ matrix polynomials in lag operator L and ε_t is a zero mean, serially uncorrelated innovations process of dimension h. In general, the components of ε_t will be correlated, such that $\mathrm{E}(\varepsilon_t\varepsilon_t') = \Sigma$. For policy analysis purposes, Σ is often decomposed as $\Sigma = \Sigma^{1/2}(\Sigma^{1/2})'$, so that $\varepsilon_t = \Sigma^{1/2}v_t$ and $\mathrm{var}(v_t) = I$. The ε's are called reduced-form innovations. The v's are referred to as (uncorrelated)

[1] The use of VARs for the purpose of policy analysis is, for example, discussed in Sims (1982, 1986) and Sargent (1984).

[2] See section 2.3.

structural or orthogonalized innovations and commonly interpreted as specific shocks to the economic system. The question of how these shocks affect the various endogenous variables in z_t is at the heart of empirical macroeconomic analyses.

It would be more realistic to assume that the parameters in (1) change over time. Ignoring such time variations represents a further potential source of misspecification, in addition to the ones considered here. However, certain types of parameter variations, such as the deterministic periodic variations considered in Tiao and Grupe (1980), can be treated as special cases of (1).

2.1. Representations of ARMA subprocesses

Let vector $z_{1,t}$, a subset of $m \leq h$ components of z_t, contain the variables that are included in the VAR model, and let z_t be ordered such that $z_t = (z'_{1,t}\, z'_{2,t})'$, where vector $z_{2,t}$ contains the $n = h - m$ variables excluded from the VAR. We refer to $\{z_{1,t}\}$ and $\{z_{2,t}\}$ as subprocesses of $\{z_t\}$. Then (1) can be partitioned such that

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$= \begin{bmatrix} B_1(L) \\ B_2(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \tag{1'}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

The moving average (MA) representation of (1) is $z_t = A^{-1}(L)B(L)\varepsilon_t = C(L)\varepsilon_t$, which, following (1′), can also be partitioned as

$$\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \begin{bmatrix} C_1(L) \\ C_2(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}. \tag{2}$$

Throughout the paper we refer to submatrices $A_{hl,i}, B_{hl,i}, B_{h,i}, C_{hl,i}, C_{h,i}$, and Σ_{hl} $(h, l = 1, 2)$ whose definitions are implied by partitionings (1′) and (2).

Partitioning (1′) gives rise to two alternative ways of representing the process of included variables, $z_{1,t}$. The first is of the form

$$A_{11}(L)z_{1,t} = -A_{12}(L)z_{2,t} + B_1(L)\Sigma^{1/2}v_t. \tag{3}$$

This representation has the advantage of including both the complete original innovations vector, ε_t, and the subset of variables excluded from z_t, $z_{2,t}$. Therefore, using (3) we can investigate the role of different types of misspecifications in terms of the complete vector z_t and the original orthogonalized innovations v_t.

An alternative way of deriving a representation for $z_{1,t}$ has been used by Quenouille (1957), Zellner and Palm (1974), Palm (1977), and others. This second representation is obtained from the partitioned ARMA representation (1′) by substituting $z_{2,t} = -A_{22}^{-1}(L)A_{21}(L)z_{1,t} + A_{22}^{-1}(L)B_2(L)\Sigma^{1/2}v_t$ into (3), to yield $\tilde{A}_1(L)z_{1,t} = \tilde{B}_1(L)v_t$, where $\tilde{A}_1(L) = a_{22}(L)A_{11}(L) - A_{12}(L)A_{22}^{+}(L)A_{21}(L)$ and $\tilde{B}_1(L) = [a_{22}(L)B_1(L) - A_{12}(L)A_{22}^{+}(L) \times B_2(L)]\Sigma^{1/2}$, with $A_{22}^{+}(L)$ and $a_{22}(L) = |A_{22}(L)|$ denoting the adjoint matrix and determinant of $A_{22}(L)$, respectively. Granger and Morris (1976) show that the sum of independent MA processes is also an MA process. Therefore, because $\tilde{B}_1(L)v_t = \tilde{B}_{11}(L)v_{1,t} + \tilde{B}_{12}(L)v_{2,t}$, the h-dimensional MA process $\tilde{B}_1(L)v_t$ can be expressed in terms of an m-dimensional MA process. Thus, $z_{1,t}$ has a proper ARMA representation of the form

$$\tilde{A}_1(L)z_{1,t} = \tilde{\tilde{B}}_1(L)u_t, \tag{4}$$

where $\tilde{A}_1(L)$ and $\tilde{\tilde{B}}_1(L)$ are $m \times m$ matrix polynomials of finite – but typically very high – degrees and u_t is a white noise process of dimension m.[3]

In principle, one could derive representation (4) from (1) and use it to analyze the effects of model misspecifications. However, (4) excludes relevant innovations, so that inconsistencies due to omitted variables cannot be analyzed directly. Furthermore, although representations (3) and (4) produce the same autocovariance function for process $z_{1,t}$, the m-dimensional innovations process u_t in (4) has lost its structural interpretation, because the h innovations v_t are compounded and reduced to the $m < h$ innovations u_t, which, in general, do not have a meaningful economic interpretation.

Our analysis below is based on representation (3), which allows us to investigate directly misspecification effects due to omitting variables $z_{2,t}$, omitting innovations, $\varepsilon_{2,t}$, and neglecting MA components. Moreover, because representation (3) retains the original set of innovations, it enables us to examine the effects of misidentifying contemporaneous relationships between the reduced-form innovations, ε_t, and the orthogonalized innovations, v_t.

[3] Provided that $\tilde{\tilde{B}}_1(L)$ in (4) is invertible, $z_{1,t}$ has the autoregressive (AR) representation $\tilde{\tilde{B}}_1^{-1}(L)\tilde{A}_1(L)z_{1,t} = u_t$. Hence, high-order AR approximations of subsets of ARMA processes can be useful in forecasting applications.

2.2. Restrictions implied by VAR specifications

If $z_{1,t}$ is generated by (1′), fitting a kth-order VAR to $z_{1,t}$ amounts, in the framework of (3), to imposing some or all of the following restrictions:

$$A_{11,i} = \mathbf{0}, \qquad i = k+1, k+2, \ldots, p, \tag{5a}$$

$$A_{12,i} = \mathbf{0}, \qquad i = 1, 2, \ldots, p, \tag{5b}$$

$$B_{11,i} = \mathbf{0}, \qquad i = 1, 2, \ldots, q, \tag{5c}$$

$$B_{12,i} = \mathbf{0}, \qquad i = 1, 2, \ldots, q. \tag{5d}$$

If the chosen lag length is too low, i.e., $k < p$, we impose restrictions (5a). If $z_{1,t}$ does not coincide with z_t, i.e., $m < h$, we impose (5b). Ignoring the MA structure of the disturbances associated with the included variables corresponds to restrictions (5c). Finally, ignoring the innovations associated with the excluded variables, $\varepsilon_{2,t}$, is captured by restrictions (5d).

Imposing any of the restrictions in (5) when estimating a VAR model for $z_{1,t}$ leads, in general, to inconsistent VAR-coefficient estimates and affects policy analyses based on derived impulse responses and variance decompositions.

2.3. Effects of misspecifications on parameter estimates

Collect the parameters of process (1) in matrix $\Theta = (A_1 \ldots A_p B_1 \ldots B_q \Sigma)$ and let $\hat{\Theta} = (\hat{A}_1 \ldots \hat{A}_p \hat{B}_1 \ldots \hat{B}_q \hat{\Sigma})$ denote the parameter estimates associated with the OLS-estimated VAR model with restrictions (5) imposed. Assuming $k \leq p$,[4] restrictions (5) imply that[5]

$$\hat{A}_i = \begin{cases} \begin{bmatrix} \hat{A}_{11,i} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \end{bmatrix}, & i = 1, 2, \ldots, k, \\ \mathbf{0}_{h\times h}, & i = k+1, k+2, \ldots, p; \end{cases} \tag{6a}$$

$$\hat{B}_i = \mathbf{0}_{h\times h}, \qquad i = 1, 2, \ldots, q; \tag{6b}$$

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{11} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \end{bmatrix}. \tag{6c}$$

[4] The assumption $k \leq p$ is for notational convenience only. If $k > p$ all results can be applied by replacing p with $p' = k$ and setting $A_i = \mathbf{0}$ for $i = p+1, p+2, \ldots, p'$.

[5] Note that it is not necessary to set the (2, 1) and (2, 2) blocks in $\hat{A}_i$ and $\hat{\Sigma}$ equal to zero, because $\tilde{A}_{11}$ in Proposition 1 does not depend on the parameters of the subprocess for $z_{2,t}$. We adopt the convention to set all parameters to zero that are ignored by representation (3). This is consistent with the framework used in section 3 and reduces the notational burden. However, setting the off-diagonal blocks in $\hat{\Sigma}$ to zero is essential, because it implies that in a VAR model the number of endogenous variables equals the number of innovations.

Let the $m \times km$ matrix $\hat{A}_{11} = (\hat{A}_{11,1}\hat{A}_{11,2} \dots \hat{A}_{11,k})$ represent the OLS estimates of the VAR parameters contained in (6a), that is,

$$\hat{A}_{11} = Z_1 X_1'(X_1 X_1')^{-1}, \tag{7}$$

where $Z_1 = (z_{1,1}z_{1,2} \dots z_{1,T})$ has dimension $m \times T$ and

$$X_1 = \begin{bmatrix} z_{1,0} & z_{1,1} & \cdots & z_{1,T-1} \\ z_{1,-1} & z_{1,0} & & z_{1,T-2} \\ \vdots & & & \vdots \\ z_{1,-k+1} & z_{1,-k+2} & \cdots & z_{1,T-k} \end{bmatrix}$$

has dimension $km \times T$.

To state the misspecification effects on the estimated VAR parameters contained in $\hat{\Theta}$, we adopt the following notation. We use tildes to denote the inconsistency of an estimator. Thus, $\tilde{\Theta} = (\tilde{A}_1 \dots \tilde{A}_p \tilde{B}_1 \dots \tilde{B}_q \tilde{\Sigma})$ denotes the inconsistency of $\hat{\Theta}$, i.e., $\text{plim}\hat{\Theta}_T = \Theta + \tilde{\Theta}$ or $\text{plim}\hat{A}_{i,T} = A_i + \tilde{A}_i$, $\text{plim}\hat{B}_{i,T} = B_i + \tilde{B}_i$, and $\text{plim}\hat{\Sigma}_T = \Sigma + \tilde{\Sigma}$, where subscript T refers to an estimator based on sample size T.

It is clear from the way (6) is partitioned that the only unknown quantities in $\tilde{\Theta}$ are $\tilde{A}_{11} = (\tilde{A}_{11,1}\tilde{A}_{11,2} \dots \tilde{A}_{11,k})$ and $\tilde{\Sigma}_{11}$. To derive these, we define the following quantities. Let

$$X_2^1 = \begin{bmatrix} z_{2,0} & z_{2,1} & \cdots & z_{2,T-1} \\ z_{2,-1} & z_{2,0} & & z_{2,T-2} \\ \vdots & & & \vdots \\ z_{2,-k+1} & z_{2,-k+2} & \cdots & z_{2,T-k} \end{bmatrix},$$

$$X_2^2 = \begin{bmatrix} z_{-k} & z_{-k+1} & \cdots & z_{T-k-1} \\ z_{-k-1} & z_{-k} & \cdots & z_{T-k-2} \\ \vdots & & & \vdots \\ z_{-p+1} & z_{-p+2} & \cdots & z_{T-p} \end{bmatrix},$$

with dimensions $kn \times T$ and $(p-k)h \times T$, respectively, and let X_2 be the matrix that contains the regressors that are omitted in the VAR estimation, i.e.,

$$X_2 = \begin{cases} X_2^1 & (kn \times T) & \text{if} \quad k = p, \\ \begin{bmatrix} X_2^1 \\ X_2^2 \end{bmatrix} & ([kn + (p-k)h] \times T) & \text{if} \quad k < p. \end{cases}$$

Collecting the coefficient matrices of polynomial $A_{12}(L)$ in (1′) in the $m \times kn$ matrix $A_{12} = (A_{12,1} A_{12,2} \dots A_{12,k})$, the autoregressive parameters associated with the omitted regressors are given by

$$A^* = \begin{cases} (A_{k+1} \dots A_p) & (h \times h(p-k)) \quad \text{if} \quad k < p,\ m = h, \\ A_{12} & (m \times kn) \quad \text{if} \quad k = p,\ m < h, \\ (A_{12} A_{11,k+1} A_{12,k+1} \dots A_{11,p} A_{12,p}) & (m \times [kn + h(p-k)]) \quad \text{if} \quad k < p,\ m < h. \end{cases} \tag{8}$$

The parameters associated with the omitted MA part are

$$B^* = (B_{1,1}\ B_{1,2} \dots B_{1,q}) \quad (m \times qn), \tag{9}$$

where submatrices $B_{1,i}$ are the coefficient matrices of polynomial $B_1(L)$ in (1′). Next, construct matrix C^* from the coefficient matrices of polynomial $C_1(L)$ in (2),

$$C^* = \begin{cases} T_C & (hq \times mq) \quad \text{if} \quad k \le q, \\ [T_C \quad \mathbf{0}_{hp \times m(k-p)}] & (hq \times mk) \quad \text{if} \quad k > q, \end{cases} \tag{10}$$

where

$$T_C = \begin{bmatrix} C'_{1,0} & \mathbf{0} & \cdots & \mathbf{0} \\ C'_{1,1} & C'_{1,0} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ C'_{1,q-1} & C'_{1,q-2} & \cdots & C'_{1,0} \end{bmatrix} \quad (hq \times mq).$$

Finally, define

$$\Gamma_{11} = \operatorname{plim} T^{-1} X_1 X'_1, \tag{11}$$

$$\Gamma_{21} = \operatorname{plim} T^{-1} X_2 X'_1, \tag{12}$$

$$\Lambda_{11,0} = \operatorname{plim} T^{-1} Z_1 Z'_1, \tag{13}$$

$$\underline{\Lambda}_1 = \operatorname{plim} T^{-1} Z_1 X'_1. \tag{14}$$

Proposition 1.[6] *The VAR-restricted OLS estimator $\hat{\Theta}$ for the subprocess $\{z_{1,t}\}$ generated by (1) is inconsistent, i.e.,* $\operatorname{plim} \hat{\Theta}_T = \Theta + \tilde{\Theta}$. *Analogous to (6), the only*

[6] The proofs of this and subsequent propositions are in the appendix.

unknown quantities in $\tilde{\Theta}$ *are* $\tilde{A}_{11} = (\tilde{A}_{11,1}\tilde{A}_{11,2} \dots \tilde{A}_{11,k})$ *and* $\tilde{\Sigma}_{11}$. *They are given by*

$$\tilde{A}_{11} = [A^*\Gamma_{21} + B^*(I_q \otimes \Sigma)C^*]\Gamma_{11}^{-1}, \tag{15}$$

with matrices $A^*, B^*, C^*, \Gamma_{11}$, *and* Γ_{21} *being defined by (8)–(12), and*

$$\tilde{\Sigma}_{11} = \Lambda_{11,0} - \underline{\Lambda}_1 \Gamma_{11}^{-1} \underline{\Lambda}_1' - \Sigma_{11}, \tag{16}$$

where $\Lambda_{11,0}$ *and* $\underline{\Lambda}_1$ *are defined by (13) and (14), respectively.*

The following corollary demonstrates how $\tilde{A}_{11}$ in (15) can be separated into additive components reflecting the different sources of misspecifications.

Corollary 1. (i) *Suppose all restrictions in (5) except (5a) are correct, i.e., the specified autoregressive order is too low. Then, inconsistency* $\tilde{A}_{11}$ *is given by* $\tilde{A}_{11} = A^*\Gamma_{21}\Gamma_{11}^{-1}$ *with* A^* *defined by the first case in (8).*

(ii) *If the inconsistency is solely due to incorrectly imposing restriction (5b), i.e., specifying a VAR model that excludes relevant variables, then* $\tilde{A}_{11} = A^*\Gamma_{21}\Gamma_{11}^{-1}$ *with* A^* *given by the second case in (8).*

(iii) *In case both (5a) and (5b) are invalid restrictions,* $\tilde{A}_{11} = A^*\Gamma_{21}\Gamma_{11}^{-1}$ *with* A^* *defined by the third case in (8).*

(iv) *If all but restrictions (5c) and (5d) are correct, i.e., neglecting the presence of MA terms, then* $\tilde{A}_{11} = B^*(I_q \otimes \Sigma)C^*\Gamma_{11}^{-1}$.

Note that $\tilde{A}_{11}$ and $\tilde{\Sigma}_{11}$ are expressed in terms of $C_{1,i}, \Gamma_{11}, \Gamma_{21}, \Lambda_{11,0}$, and $\underline{\Lambda}_1$. The $C_{1,i}$'s are the coefficient matrices of polynomial $C_1(L)$ defined in (2), and matrices $\Gamma_{11}, \Gamma_{21}, \Lambda_{11,0}$, and $\underline{\Lambda}_1$ consist of (submatrices of) the theoretical autocovariances $\Lambda_k = \mathrm{cov}(z_t z_{t-k}')$. Combining the closed-form expression for theoretical autocovariances derived in Mittnik (1990) with (15) and (16) enables us to relate inconsistencies $\tilde{A}_{11}$ and $\tilde{\Sigma}_{11}$ directly to the parameters of the underlying ARMA process (1).

2.4. Examples of misspecification effects

We use three simple examples to illustrate the results stated in Proposition 1 and Corollary 1. Specifying three different weakly stationary processes representing the underlying system (1) and assuming, in each case, a modeler estimates the univariate AR(1) model $y_t = \alpha y_{t-1} + v_t$, we derive $\operatorname{plim} \hat{A}_{11,T} = A_{11} + \tilde{A}_{11}$ or, in the notation of the examples, $\operatorname{plim} \hat{\alpha}_T = a_1 + \tilde{\alpha}$.

First, let (1) be the univariate AR(2) process

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t, \tag{17}$$

and assume the analyst fits the AR(1) model $y_t = \alpha y_{t-1} + v_t$. Then, the OLS estimator of α, $\hat{\alpha}$, corresponds to $\hat{A}_{11}$ in (7) and its inconsistency, $\tilde{\alpha}$, corresponds to $\tilde{A}_{11}$ in (15). Given that the specified lag length is too low, it follows from case (i) of Corollary 1 that plim $\hat{\alpha}_T = a_1 + a_2\lambda_1/\lambda_0$, where λ_0 and λ_1 are the theoretical variance and first-order autocovariance of process (17), or, because the first-order autocorrelation $\lambda_1/\lambda_0 = a_1/(1 - a_2)$,

$$\operatorname{plim} \alpha_T = a_1 + a_2 \frac{a_1}{1 - a_2}. \tag{18}$$

Next, let (1) be the bivariate AR(1) process

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}, \tag{19}$$

and suppose the univariate AR(1) model $y_{1,t} = \alpha y_{1,t-1} + v_t$ is estimated for variable $y_{1,t}$. Then, as stated in case (ii) of Corollary 1, the exclusion of variable $y_{2,t-1}$ implies that

$$\operatorname{plim} \hat{\alpha}_T = a_{11} + a_{12} \frac{\lambda_{0,12}}{\lambda_{0,11}}, \tag{20}$$

where $\lambda_{0,11}$ and $\lambda_{0,12}$ are the (1,1) and (1,2) elements of Λ_0, the theoretical covariance matrix of the underlying bivariate process (19).

Finally, let (1) be the univariate ARMA(1,1) process

$$y_t = a_1 y_{t-1} + \varepsilon_t + b_1 \varepsilon_{t-1}, \qquad \operatorname{var}(\varepsilon_t) = \sigma^2, \tag{21}$$

and assume the modeler estimates the AR(1) model $y_t = \alpha y_{t-1} + v_t$. Then, as stated in case (iii) of Corollary 1, omitting the MA component implies that

$$\operatorname{plim} \hat{\alpha}_T = a_1 + \frac{b_1 \sigma^2 c_0}{\lambda_0} = a_1 + \frac{(1 - a_1)^2 b_1}{1 + 2a_1 b_1 + b_1^2}, \tag{22}$$

where $c_0 = 1$ is the initial coefficient of the pure moving average representation of (21), i.e., $y_t = \sum_{i=0}^{\infty} c_i \varepsilon_{t-i}$, and $\lambda_0 = \sigma^2(1 + 2a_1 b_1 + b_1^2)/(1 - a_1^2)$ is the theoretical variance of process (21). By setting the AR coefficient a_1 in (21) and (22) to zero, the last example also illustrates that fitting an AR(1) model to data generated by an MA(1) process implies that plim $\hat{\alpha}_T = \tilde{\alpha} = b_1/(1 + b_1^2)$.

3. Misspecifications and innovation accounting

In this section we use the results in Proposition 1 and derive the effects of VAR misspecifications on impulse responses and variance decompositions. To avoid repetitious definitions and to clearly distinguish between matrices constructed from different quantities, we use arguments Θ, $\operatorname{plim}(\hat{\Theta}_T)$, or $\tilde{\Theta}$ to indicate whether a matrix is constructed or derived from actual parameter values, plim values, or inconsistencies.

3.1. Moving average coefficients and impulse responses

The MA representation (MAR) of (1) is $z_t = C(L)\varepsilon_t$, where $C(L) = A^{-1}(L)B(L) = I + C_1 L + C_2 L^2 + \cdots$ with C_i representing the ith MAR-coefficient matrix. The (orthogonalized) impulse responses of (1), denoted by Ψ_i, are related to the MAR coefficients by $\Psi_i = C_i \Sigma^{1/2}$.

Assume the modeler is interested in investigating the first $r + 1$ impulse responses, where r is a nonnegative integer. Let the true MAR coefficients, $C_i(\Theta)$, and impulse responses, $\Psi_i(\Theta)$, for $i = 0, \ldots, r$, of the underlying ARMA process (1) be collected in $m(r+1) \times m$ matrices $C(\Theta) = [C'_0(\Theta) \ldots C'_r(\Theta)]'$ and $\Psi(\Theta) = [\Psi'_0(\Theta) \ldots \Psi'_r(\Theta)]'$, respectively. Then, following Mittnik (1987), the MAR coefficients and impulse responses of (1) are obtained by

$$C(\Theta) = M^{-1}(\Theta)B(\Theta), \tag{23}$$

$$\Psi(\Theta) = M^{-1}(\Theta)B(\Theta)\Sigma^{1/2}(\Theta), \tag{24}$$

where $B(\Theta) = [B'_0(\Theta) \ldots B'_r(\Theta)]'$ is $m(r+1) \times m$ and $M(\Theta)$ is an $m(r+1) \times m(r+1)$ lower-triangular block Toeplitz matrix whose first block column is given by $[I, -A'_1(\Theta), \ldots, -A'_r(\Theta)]'$.[7] We use $C(\operatorname{plim}\hat{\Theta}_T)$ and $\Psi(\operatorname{plim}\hat{\Theta}_T)$ to denote MAR coefficients and impulse responses obtained from $\operatorname{plim}\hat{\Theta}_T$, where $\hat{\Theta}_T$ is the VAR-restricted estimator implied by (6) and (7), and use Δ with the appropriate subscript to indicate the inconsistency of quantities derived from $\operatorname{plim}\hat{\Theta}_T$. For example, $\operatorname{plim}\Psi(\hat{\Theta}_T) = \Psi(\operatorname{plim}\hat{\Theta}_T) = \Psi(\Theta + \tilde{\Theta}) = \Psi(\Theta) + \Delta_\Psi$, where $\Delta_\Psi = [\Delta'_{\Psi_0} \ldots \Delta'_{\Psi_r}]'$.

Proposition 2. MAR coefficients and impulse responses derived from the VAR-restricted estimator $\hat{\Theta}$ are inconsistent, i.e., $\operatorname{plim} C(\hat{\Theta}_T) = C(\Theta) + \Delta_C$ and $\operatorname{plim}\Psi(\hat{\Theta}_T) = \Psi(\Theta) + \Delta_\Psi$, where

$$\Delta_C = M^{-1}(\operatorname{plim}\hat{\Theta}_T)B(\tilde{\Theta}) + N(\tilde{\Theta})B(\Theta), \tag{25}$$

$$\Delta_\Psi = \Delta_C \Sigma^{1/2}(\Theta) + M^{-1}(\operatorname{plim}\hat{\Theta}_T)B(\operatorname{plim}\hat{\Theta}_T)\tilde{\Sigma}^{1/2}, \tag{26}$$

with $N(\tilde{\Theta}) = -M^{-1}(\Theta)M(\tilde{\Theta})[I + M^{-1}(\Theta)M(\tilde{\Theta})]^{-1}M^{-1}(\Theta)$.

[7] It is implicitly assumed that $A_i = \mathbf{0}$ for $i = p+1, \ldots, r$, if $r > p$, and $B_i = \mathbf{0}$ for $i = q+1, \ldots, r$, if $r > q$.

As is clear from (26), an additional source of misspecification, namely the decomposition of $\Sigma(\hat{\Theta})$, has to be considered when deriving the inconsistency for the orthogonalized impulse responses. This gives rise to:

Corollary 2. Suppose that the restrictions in (5) are valid, but the chosen decomposition of the residual covariance matrix is incorrect, then, because $\Delta_C = \mathbf{0}$ *and* $\operatorname{plim} \hat{\Theta}_T = \Theta$, *the inconsistency of the estimated impulse responses is* $\Delta_\Psi = M^{-1}(\Theta)B(\Theta)\tilde{\Sigma}^{1/2}$.

3.2. Variance decompositions

Let $D_\tau(\Theta)$ denote the $h \times h$ variance decomposition matrix for forecast horizon τ that corresponds to the underlying ARMA process (1). Entry (i,j) in $D_\tau(\Theta)$ represents the share of the expected τ-step-ahead squared prediction error for variable i which is due to innovation j. $D_\tau(\Theta)$ is computed by $D_\tau(\Theta) = W_\tau(\Theta)\sum_{k=0}^{\tau-1}\Psi_k^{(2)}(\Theta)$, where $X^{(2)}$ denotes the Hadamard product [see, for example, Magnus and Neudecker (1989, p. 45)] of a matrix with itself, i.e., $X \odot X = X^{(2)} = (x_{ij}^2)$. The diagonal $h \times h$ matrix W_τ contains the inverses of the τ-step-ahead prediction-error variances for the variables in z_t, i.e., $W_\tau(\Theta) = \operatorname{diag}^{-1}[\sum_{k=0}^{\tau-1}\Psi_k^{(2)}(\Theta)\mathbf{1}_h]$, where $\mathbf{1}_h$ denotes an $h \times 1$ vector of ones.

Proposition 3. The inconsistency of $D_\tau(\hat{\Theta})$, *the* τ*-horizon variance decomposition matrix derived from the VAR-restricted estimator* $\hat{\Theta}$, $\Delta_{D_\tau} = \operatorname{plim} D_\tau(\hat{\Theta}_T) - D_\tau(\Theta)$, *is*

$$\Delta_{D_\tau} = W_\tau(\Theta)\sum_{k=0}^{\tau-1}[2\Psi_k(\Theta) + \Delta_{\Psi_k}] \odot \Delta_{\Psi_k} + \Delta_{W_\tau}\sum_{l=0}^{\tau-1}\Psi_l^{(2)}(\operatorname{plim}\hat{\Theta}_T), \quad (27)$$

where

$$\Delta_{W_\tau} = \begin{bmatrix} \operatorname{diag}^{-1}\left[\sum_{k=0}^{\tau-1}\Psi_{11,k}^{(2)}(\operatorname{plim}\hat{\Theta}_T)\mathbf{1}_m\right] & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - W_\tau(\Theta).$$

Corollary 3. Suppose that restrictions (5) are valid, but the chosen decomposition of the estimated residual covariance matrix is incorrect. Then, because $\operatorname{plim} W_\tau(\hat{\Theta}_T)$ *is not affected by the chosen decomposition, i.e.,* $\Delta_{W_\tau} = \mathbf{0}$, *the inconsistency of the estimated variance decomposition is given by* $\Delta_{D_\tau} = W_\tau(\Theta)\sum_{k=0}^{\tau-1}[2\Psi_k(\Theta) + \Delta_{\Psi_k}] \odot \Delta_{\Psi_k}$, *where the* Δ_{Ψ_k}*'s are as in Corollary 2.*

4. Quantitative illustration

In this section we employ the analytical results of sections 2 and 3 to illustrate the role misspecifications may play in applied work. Using aggregate, quarterly, postwar U.S. time series data on investment expenditures, the price of investment, and a discount rate, we estimate a trivariate ARMA model.[8] According to Akaike's information criterion (AIC), an ARMA(1, 1) model was most appropriate for the trivariate system. The impulse responses implied by this estimated ARMA(1, 1) model are plotted in fig. 1. The predicted error variances and variance decompositions are presented in tables 1 and 2, respectively.

Treating the estimated ARMA(1, 1) model as the true, underlying economic system, we derive the inconsistencies associated with various misspecified VARs. We separately investigate each of the potential sources of inconsistency discussed in section 2 – i.e., omitted MA components, omitted variables, and misspecified lag lengths. The results in sections 2 and 3 allow us to determine the inconsistencies of the implied impulse responses, prediction-error variances,[9] and variance decompositions directly from the original ARMA(1, 1) parameters.

The first experiment we conduct quantifies effects of ignoring the MA part of the original ARMA(1, 1) process, i.e., we assume the modeler fits a trivariate VAR(1) model. The impulse responses of the implied VAR(1) model are shown in fig. 1, together with the actual ARMA(1, 1) responses. The ordering of the variables used in the Cholesky decomposition corresponds to the order of appearance in fig. 1.

Inspecting the implied VAR(1) responses, we see that omitting the MA component causes dramatic changes in the response of investment expenditures to a shock in the discount rate, the response of the deflator to a shock in both investment expenditures and the discount rate, as well as the response of the discount rate to a shock in investment expenditures. Taking parameter uncertainty into account, we find that only two of the nine original ARMA(1, 1) responses – namely, the response of investment to a deflator shock and the deflator response to a discount rate shock – are completely within the two-standard-error bounds.[10] All other responses deviate significantly from those of the ARMA(1, 1) process for one or more periods. Except for all one-step-ahead and one of the two-step-ahead ones, the prediction-error variances (table 1) do not deviate significantly from those of the ARMA(1, 1) process. The primary

[8] The data we use are the National Income and Product Accounts' real gross fixed private investment series, its deflator, and a short-term commercial paper rate from the Citibase data bank. Because the investment and deflator series are nonstationary, we took their first difference in logs. This amounts to assuming that the variables are I(1), but not cointegrated.

[9] As follows from section 3.2, the τ-step prediction-error variances are computed from $\sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\Theta)\mathbf{1}_h$.

[10] The standard errors were generated by a Monte Carlo simulation based on 400 replications.

Table 1

Prediction-error variances (the prediction-error variances and their standard errors, given in parentheses, are multiplied by 10^3).

Periods ahead	Investment	Deflator	Discount rate
	(A) *Three-variable ARMA(1,1) model*		
1	0.28	0.05	0.05
2	0.77	0.06	0.17
4	1.19	0.09	0.36
8	1.37	0.11	0.54
12	1.41	0.13	0.62
24	1.45	0.14	0.69
	(B) *Three-variable VAR(1) model*		
1	0.70 (0.09)	0.06 (0.01)	0.08 (0.01)
2	0.93 (0.14)	0.10 (0.02)	0.15 (0.02)
4	1.11 (0.20)	0.13 (0.03)	0.28 (0.05)
8	1.28 (0.28)	0.16 (0.05)	0.46 (0.12)
12	1.37 (0.33)	0.17 (0.07)	0.56 (0.18)
24	1.46 (0.40)	0.19 (0.09)	0.69 (0.30)
	(C) *Three-variable VAR(6) model*		
1	0.31 (0.04)	0.04 (0.06)	0.05 (0.01)
2	0.76 (0.11)	0.07 (0.01)	0.17 (0.03)
4	1.18 (0.23)	0.11 (0.04)	0.35 (0.08)
8	1.41 (0.33)	0.18 (0.08)	0.56 (0.18)
12	1.50 (0.36)	0.22 (0.10)	0.67 (0.25)
24	1.60 (0.40)	0.26 (0.13)	0.80 (0.32)
	(D) *Two-variable VAR(6) model*		
1	0.73 (0.08)	0.06 (0.01)	NA
2	1.00 (0.13)	0.15 (0.02)	NA
4	1.26 (0.19)	0.25 (0.04)	NA
8	1.40 (0.23)	0.32 (0.08)	NA
12	1.44 (0.25)	0.33 (0.11)	NA
24	1.48 (0.28)	0.34 (0.14)	NA
	(E) *Two-variable VAR(2) model*		
1	0.73 (0.09)	0.06 (0.01)	NA
2	0.99 (0.15)	0.08 (0.01)	NA
4	1.21 (0.22)	0.12 (0.02)	NA
8	1.35 (0.29)	0.15 (0.04)	NA
12	1.39 (0.32)	0.16 (0.05)	NA
24	1.42 (0.34)	0.17 (0.06)	NA

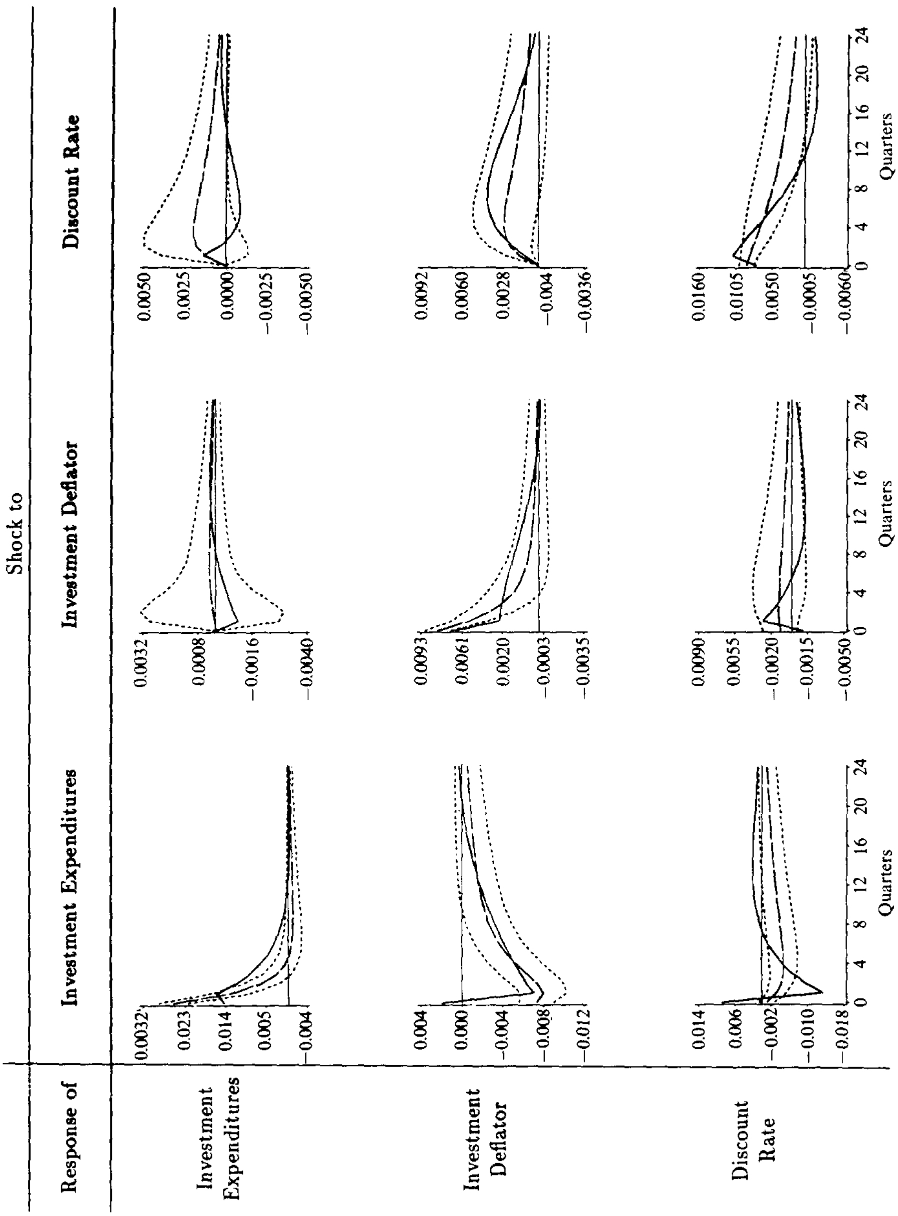

Fig. 1. Impulse response functions – Trivariate VAR(1). Key: trivariate ARMA (1,1) ——, trivariate VAR(1) – – –, ± 2 std. deviation - - - -.

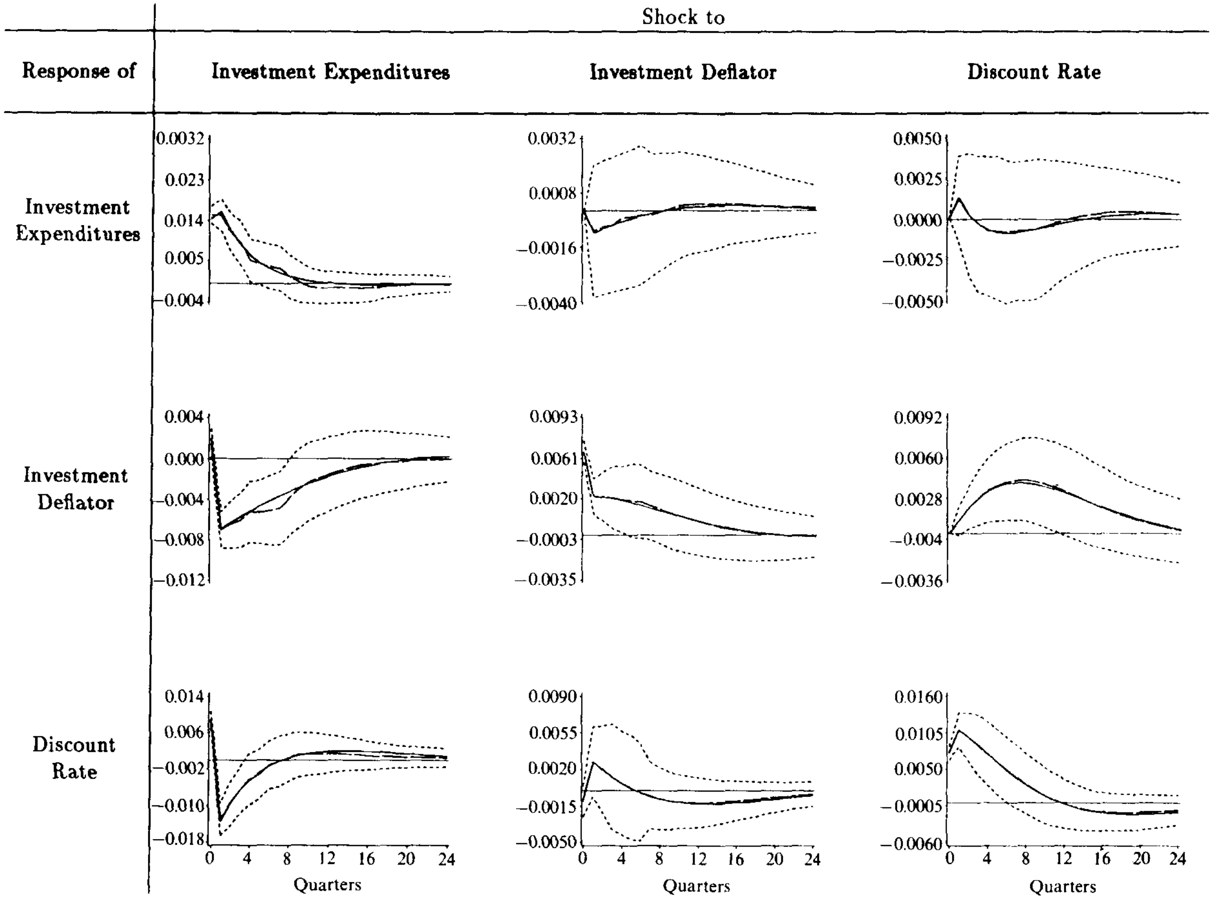

Fig. 2. Impulse response functions – Trivariate VAR(6). Key: trivariate ARMA (1,1) —— , trivariate VAR(6) – – –, ± 2 std. deviation - - - -.

Table 2

Variance decompositions (standard errors are shown in parentheses).

(A) Variance decomposition of investment expenditures

Model	Periods ahead	Proportions due to: Investment	Deflator	Discount rate
Three-variable ARMA(1, 1)	1	0.71	0.01	0.27
	2	0.59	0.07	0.34
	4	0.55	0.11	0.34
	8	0.53	0.15	0.31
	12	0.52	0.17	0.31
	24	0.51	0.17	0.32
Three-variable VAR(1)	1	0.92 (0.01)	0.08 (0.01)	0.00 (0.00)
	2	0.86 (0.02)	0.13 (0.02)	0.01 (0.01)
	4	0.76 (0.04)	0.19 (0.03)	0.05 (0.03)
	8	0.67 (0.07)	0.22 (0.04)	0.11 (0.05)
	12	0.63 (0.08)	0.23 (0.04)	0.14 (0.07)
	24	0.60 (0.10)	0.22 (0.14)	0.17 (0.08)
Three-variable VAR(6)	1	0.74 (0.03)	0.01 (0.01)	0.25 (0.03)
	2	0.60 (0.05)	0.07 (0.01)	0.33 (0.04)
	4	0.56 (0.05)	0.11 (0.02)	0.33 (0.05)
	8	0.52 (0.06)	0.17 (0.03)	0.30 (0.05)
	12	0.51 (0.06)	0.19 (0.04)	0.30 (0.05)
	24	0.51 (0.06)	0.19 (0.05)	0.30 (0.15)
Two-variable VAR(6)	1	0.89 (0.02)	0.11 (0.02)	NA
	2	0.79 (0.02)	0.21 (0.02)	NA
	4	0.71 (0.03)	0.29 (0.03)	NA
	8	0.65 (0.05)	0.35 (0.05)	NA
	12	0.64 (0.06)	0.36 (0.06)	NA
	24	0.64 (0.06)	0.36 (0.06)	NA
Two-variable VAR(2)	1	0.89 (0.02)	0.11 (0.02)	NA
	2	0.81 (0.02)	0.19 (0.02)	NA
	4	0.74 (0.03)	0.26 (0.03)	NA
	8	0.67 (0.05)	0.33 (0.05)	NA
	12	0.65 (0.06)	0.35 (0.06)	NA
	24	0.64 (0.07)	0.36 (0.07)	NA

(B) Variance decomposition of investment deflator

Model	Periods ahead	Proportions due to: Investment	Deflator	Discount rate
Three-variable ARMA(1, 1)	1	0.00	0.97	0.03
	2	0.02	0.85	0.14
	4	0.02	0.81	0.16
	8	0.02	0.85	0.13
	12	0.02	0.83	0.15
	24	0.02	0.75	0.23
Three-variable VAR(1)	1	0.00 (0.00)	0.98 (0.03)	0.02 (0.03)
	2	0.00 (0.03)	0.97 (0.05)	0.03 (0.04)
	4	0.00 (0.06)	0.95 (0.08)	0.05 (0.07)
	8	0.00 (0.07)	0.91 (0.11)	0.08 (0.10)
	12	0.00 (0.07)	0.89 (0.12)	0.10 (0.12)
	24	0.01 (0.07)	0.87 (0.14)	0.12 (0.14)
Three-variable VAR(6)	1	0.00 (0.00)	0.98 (0.03)	0.02 (0.04)
	2	0.01 (0.06)	0.85 (0.09)	0.13 (0.10)
	4	0.02 (0.10)	0.82 (0.12)	0.16 (0.12)

Table 2 (continued)

Model	Periods ahead	(B) Variance decomposition of investment deflator Proportions due to Investment	Deflator	Discount rate
	8	0.02 (0.09)	0.86 (0.11)	0.13 (0.11)
	12	0.02 (0.09)	0.84 (0.12)	0.14 (0.11)
	24	0.02 (0.09)	0.78 (0.12)	0.20 (0.11)
Two-variable VAR(6)	1	0.00 (0.00)	1.00 (0.02)	NA
	2	0.00 (0.07)	1.00 (0.07)	NA
	4	0.00 (0.10)	1.00 (0.10)	NA
	8	0.02 (0.13)	0.98 (0.13)	NA
	12	0.05 (0.14)	0.95 (0.14)	NA
	24	0.07 (0.14)	0.93 (0.14)	NA
Two-variable VAR(2)	1	0.00 (0.00)	1.00 (0.00)	NA
	2	0.00 (0.06)	1.00 (0.06)	NA
	4	0.00 (0.07)	1.00 (0.07)	NA
	8	0.01 (0.09)	0.99 (0.09)	NA
	12	0.01 (0.10)	0.99 (0.10)	NA
	24	0.01 (0.10)	0.99 (0.10)	NA

Model	Periods ahead	(C) Variance decomposition of discount rate Proportions due to Investment	Deflator	Discount rate
Three-variable ARMA(1, 1)	1	0.00	0.00	1.00
	2	0.01	0.01	0.98
	4	0.01	0.04	0.96
	8	0.01	0.13	0.86
	12	0.01	0.21	0.78
	24	0.01	0.25	0.74
Three-variable VAR(1)	1	0.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	2	0.01 (0.03)	0.01 (0.01)	0.97 (0.03)
	4	0.03 (0.06)	0.05 (0.03)	0.92 (0.05)
	8	0.05 (0.09)	0.09 (0.06)	0.86 (0.06)
	12	0.05 (0.09)	0.11 (0.07)	0.84 (0.06)
	24	0.06 (0.09)	0.12 (0.07)	0.83 (0.06)
Three-variable VAR(6)	1	0.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	2	0.01 (0.02)	0.01 (0.01)	0.99 (0.02)
	4	0.00 (0.04)	0.04 (0.02)	0.96 (0.04)
	8	0.01 (0.06)	0.14 (0.05)	0.86 (0.07)
	12	0.01 (0.07)	0.22 (0.08)	0.77 (0.09)
	24	0.01 (0.07)	0.27 (0.10)	0.73 (0.10)
Two-variable VAR(6)	1	NA	NA	NA
	2	NA	NA	NA
	4	NA	NA	NA
	8	NA	NA	NA
	12	NA	NA	NA
	24	NA	NA	NA
Two-variable VAR(2)	1	NA	NA	NA
	2	NA	NA	NA
	4	NA	NA	NA
	8	NA	NA	NA
	12	NA	NA	NA
	24	NA	NA	NA

effect the omission of the MA component has on the variance decompositions is that the discount rate explains less of the variation of the other variables in the system. In the decomposition of investment (table 2) the discount rate explains less than 5% of the variation in investment for quarters 1 through 4 for the VAR(1) model, compared to approximately 30% in the original ARMA(1, 1) system. The importance of discount-rate innovations in explaining variations in the investment deflator decreases similarly.

Ignoring the MA part and yet keeping the AR order fixed at one is an unrealistic experiment, given that mixed ARMA processes can be approximated by higher-order AR models. It turns out that the Bayesian information criterion (BIC) favors the VAR(1) considered above, whereas the AIC selects a VAR(6). Therefore, we derive the implied VAR(6) model, whose impulse responses are shown in fig. 2, together with those of the original ARMA(1, 1) process. The VAR(6) does well in approximating the impulse responses, prediction-error variances (table 1), and variance decompositions (table 2) of the ARMA(1, 1) process. This demonstrates that adding higher-order AR terms may be sufficient to overcome misspecification problems due to omitted MA components.

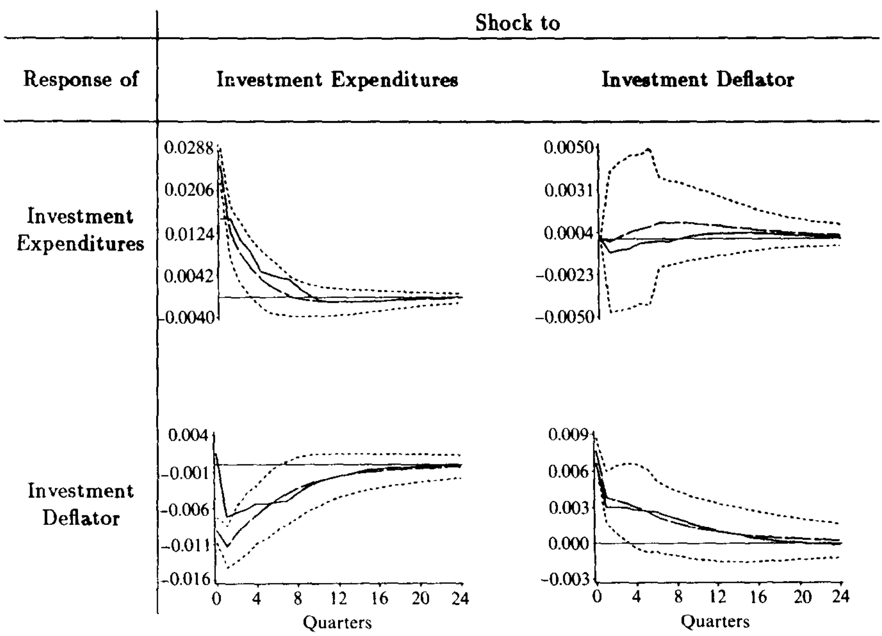

Fig. 3. Impulse response functions – Bivariate VAR(6). Key: trivariate VAR(6) ———, bivariate VAR(6) – – – –, ± 2 std. deviations - - - - - -.

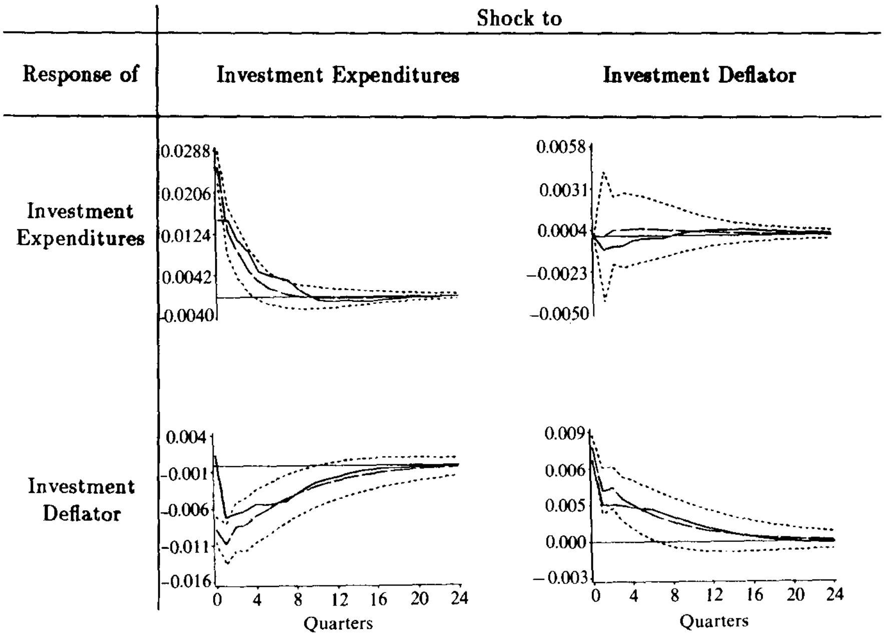

Fig. 4. Impulse response functions – Bivariate VAR(2). Key: trivariate VAR(6) ———, bivariate VAR(2) – – – –, ± 2 std. deviations - - - - - -.

The purpose of the second experiment is to consider the role of omitted variables in estimated VARs. To do this, we let the trivariate VAR(6) model, derived in the previous experiment, take the role of the true, underlying system, i.e., (1) now corresponds to a VAR(6) process. We investigate the omitted variable issue by excluding the discount rate from the trivariate VAR(6) system and derive the response functions, prediction-error variances, and variance decompositions of the implied bivariate model using the results in sections 2 and 3.

The resulting impulse responses, displayed in fig. 3 together with the actual trivariate VAR(6) responses, show that for the bivariate VAR(6) two of the four response functions – namely the investment and deflator responses to a shock in investment – deviate significantly from the trivariate VAR(6) responses, but do so only for contemporaneous or one-step responses. The prediction-error variances (table 1) differ little from those of the trivariate VAR(6). The bivariate variance decompositions (table 2), on the other hand, deviate significantly, indicating that the investment deflator absorbs almost all of the effect of the omitted discount rate.

Again, as with the previous experiment, fitting a VAR(6) model to both the trivariate and bivariate system without searching for the optimal lag length is unrealistic. Given that both the AIC and BIC favor a VAR(2) model for the bivariate process, we also derive its implied responses, prediction-error variances, and variance decompositions.

As was the case for the bivariate VAR(6) model, the impulse response functions (fig. 4), and prediction-error variances (table 1) for the bivariate VAR(2) are, for the most part, not significantly different from those of the original trivariate VAR(6). The variance decompositions (table 2), however, differ significantly from those of the trivariate VAR(6) because the deflator seems to absorb all of the variation associated with the omitted discount rate.

The first experiment, where we omit the MA part from the ARMA(1,1) process, illustrates implicitly the importance of specifying the lag length and the consequences of using alternative order-selection criteria in VAR modeling. Assuming that the true process has a trivariate VAR(6) representation, as chosen by the AIC, but the analyst fits a trivariate VAR(1), chosen by the BIC, then the inconsistencies of the resulting impulse responses, prediction-error variances, and variance decomposition correspond roughly to the inconsistencies due to omitting the MA part from the original ARMA(1,1) process.[11]

5. Conclusions

The quantitative results of this paper indicate that misspecification effects on impulse responses, prediction-error variances, and variance decompositions can be dramatic as far as their point estimates are concerned. However, the importance of the inconsistencies diminishes substantially when taking parameter uncertainty into account. While impulse responses and prediction-error variances appear to be more robust, variance decompositions, which are essentially ratios of potentially inconsistent quantities, are more sensitive to misspecifications – even when taking parameter uncertainty into account.

Moreover, we find that inconsistencies due to omitted variables and lag-length misspecification are potentially more serious, while ignoring MA components may be overcome by appropriately chosen AR orders. However, overparameterization problems may limit this strategy given the sample sizes typically encountered in macroeconomic studies.

The results of this paper are enlightening because it has been implicitly assumed by VAR modelers that, although there exist misspecification problems with omitted variables, uncertain lag lengths, and the presence of MA components, their effects are tolerable in practice. In view of the difficulties

[11] An alternative to using information criteria is to calculate posterior probabilities of alternative models. See, for example, Moulton (1991) and Steel and Richard (1991) for examples. One needs to recognize, however, that the use of order-selection criteria, such as AIC, BIC, or posterior-odds ratios, involves pretesting and, thus, affects subsequent inference.

demonstrated here, as well as the already existing criticisms in the literature, the value of the VAR approach for policy analyses has to be questioned. Other alternatives, such as the structural econometric modeling time series approach suggested in Zellner and Palm (1974) and Zellner (1987) or Bayesian approaches in the spirit of Doan et al. (1984), should be considered. At least, however, the extent to which structural VAR analyses are sensitive to misspecification errors has to be considered by the investigator.

Appendix

Proof of Proposition 1. Here we prove (15) and (16). To derive $\underline{A}_1 = \operatorname{plim} T^{-1} Z_1 X_1'$, write the T difference equations that correspond to (3) in matrix form, to express Z_1 as

$$Z_1 = A_{11} X_1 + A^* X_2 + B^* U + E_1, \tag{A.1}$$

where $E_1 = (\varepsilon_{1,1} \varepsilon_{1,2} \dots \varepsilon_{1,T})$, $A_{11} = (A_{11,1} A_{11,2} \dots A_{11,k})$, and

$$U = \begin{bmatrix} \varepsilon_{1,0} & \varepsilon_{1,1} & \cdots & \varepsilon_{1,T-1} \\ \varepsilon_{1,-1} & \varepsilon_{1,0} & & \varepsilon_{1,T-2} \\ \vdots & & & \vdots \\ \varepsilon_{1,-k+1} & \varepsilon_{1,-k+2} & \cdots & \varepsilon_{1,T-k} \end{bmatrix}.$$

To prove (15), substitute the RHS of (A.1) into the normal equations $\hat{A}_{11} X_1 X_1' = Z_1 X_1'$ and take probability limits, to obtain

$$\hat{A}_{11} \Gamma_{11} = A_{11} \Gamma_{11} + A^* \Gamma_{21} + B^* \operatorname{plim} T^{-1} U X_1'. \tag{A.2}$$

The (i, j)th (block) entry of UX_1' is of the form $\sum_{k=0}^{T-1} \varepsilon_{1-i+k} z'_{1,k-j+1}$. Replacing the $z_{1,t}$'s by the moving average representation implied by (2), i.e., $z_{1,t} = C_1(L)\varepsilon_t$, gives

$$\operatorname{plim} T^{-1} \sum_{k=0}^{T-1} \left(\varepsilon_{1-i+k} \sum_{l=0}^{\infty} \varepsilon'_{k-j+1-l} C'_{1,l} \right) = \begin{cases} \Sigma C'_{1,i-j} & \text{if} \quad i \geq j, \\ \mathbf{0} & \text{if} \quad i < j. \end{cases}$$

Therefore, $\operatorname{plim} T^{-1} U X_1' = (I_q \otimes \Sigma) C^*$. Because $\operatorname{plim} T^{-1} E_1 X_1' = \mathbf{0}$, post-multiplying (A.2) by Γ_{11}^{-1} and subtracting A_{11} gives (15).

To show (16), rewrite [analogous to (A.1)] the T equations estimated by the VAR modeler in matrix terms, i.e., $\hat{\mathscr{E}}_1 = Z_1 - \hat{A}_{11} X_1$. Replacing $\hat{A}_{11}$ by the

RHS of (7) gives $\hat{\mathscr{E}}_1 = Z_1 - Z_1 X_1'(X_1 X_1')^{-1} X_1$. Therefore, $\hat{\mathscr{E}}_1 \hat{\mathscr{E}}_1' = Z_1 Z_1' - Z_1 X_1'(X_1 X_1')^{-1} X_1 Z_1'$. Subtracting Σ_{11} from both sides of the last expression and taking probability limits yields (16). ∎

The proofs of Propositions 2 and 3 use the Slutsky Theorem [see, for example, Greenberg and Webster (1983, p. 8)], which states that if $f(X)$ is a continuous function, then $\operatorname{plim} f(X_T) = f(\operatorname{plim} X_T)$, where $f(X)$: $\mathbf{R}^{m \times n} \to \mathbf{R}^{p \times q}$ and X is constant or a random variable.

Proof of Proposition 2. $\Delta_C = \operatorname{plim} M^{-1}(\hat{\Theta}_T) B(\hat{\Theta}_T) - M^{-1}(\Theta) B(\Theta)$, from (25). Because $\operatorname{plim}(\hat{\Theta}_T) = \Theta + \tilde{\Theta}$, $\Delta_C = M^{-1}(\Theta + \tilde{\Theta}) B(\Theta + \tilde{\Theta}) - M^{-1}(\Theta) B(\Theta)$. The definitions of $M(\cdot)$ and $B(\cdot)$ imply that $M(\Theta + \tilde{\Theta})$ and $B(\Theta + \tilde{\Theta})$ can be written as $M(\Theta) + M(\tilde{\Theta})$ and $B(\Theta) + B(\tilde{\Theta})$, respectively. Therefore,

$$
\begin{aligned}
\Delta_C &= [M(\Theta) + M(\tilde{\Theta})]^{-1} [B(\Theta) + B(\tilde{\Theta})] - M^{-1}(\Theta) B(\Theta) \\
&= [M(\Theta) + M(\tilde{\Theta})]^{-1} B(\Theta) + M^{-1}(\operatorname{plim} \hat{\Theta}_T) B(\tilde{\Theta}) \\
&\quad - M^{-1}(\Theta) B(\Theta).
\end{aligned}
$$

Using $[M(\Theta) + M(\tilde{\Theta})]^{-1} = M^{-1}(\Theta) - M^{-1}(\Theta)[I + M^{-1}(\Theta) M(\tilde{\Theta})]^{-1} \times M^{-1}(\Theta) = M^{-1}(\Theta) + N(\tilde{\Theta})$, we obtain (25). The proof of (26) is analogous. ∎

Proof of Proposition 3. From the definition of $D_\tau(\Theta)$ in section 3.2 we have $\Delta_{D_\tau} = \operatorname{plim}[W_\tau(\hat{\Theta}_T) \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\hat{\Theta}_T)] - W_\tau(\Theta) \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\Theta)$. Because $\Delta_{W_\tau} = W_\tau(\operatorname{plim} \hat{\Theta}_T) - W_\tau(\Theta)$ and $\Psi_k(\operatorname{plim} \hat{\Theta}_T) = \Psi_k(\Theta) + \Delta_{\Psi_k}$,

$$
\begin{aligned}
\Delta_{D_\tau} &= [W_\tau(\Theta) + \Delta_{W_\tau}] \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\operatorname{plim} \hat{\Theta}_T) - W_\tau(\Theta) \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\Theta) \\
&= W_\tau(\Theta) \sum_{k=0}^{\tau-1} [\Psi_k(\Theta) + \Delta_{\Psi_k}]^{(2)} + \Delta_{W_\tau} \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\operatorname{plim} \hat{\Theta}_T) \\
&\quad - W_\tau(\Theta) \sum_{k=0}^{\tau-1} \Psi_k^{(2)}(\Theta).
\end{aligned}
$$

Then, (27) follows from the fact that $[\Psi_k(\Theta) + \Delta_{\Psi_k}]^{(2)} = \Psi_k^{(2)}(\Theta) + \Delta_{\Psi_k}^{(2)} + 2\Psi_k(\Theta) \odot \Delta_{\Psi_k}$. ∎

References

Doan, T., R.B. Litterman, and C.A. Sims, 1984, Forecasting and conditional projection using realistic prior distributions, Econometric Review 3, 1–100.

Greenberg, E. and C.E. Webster, 1983, Advanced econometrics: A bridge to the literature (Wiley, New York, NY).

Lütkepohl, H., 1990, Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models, Review of Economics and Statistics 72, 116–125.

Magnus, J.R. and H. Neudecker, 1988, Matrix differential calculus (Wiley, Chichester).

Mittnik, S., 1987, Nonrecursive methods for computing the coefficients of the autoregressive and the moving-average representation of mixed ARMA processes, Economics Letters 23, 279–284.

Mittnik, S., 1990, Computation of theoretical autocovariance matrices of multivariate autoregressive moving-average time series, Journal of the Royal Statistical Society B 52, 151–155.

Moulton, B.R., 1991, A Bayesian approach to regression selection and estimation, with applications to a price index for radio services, Journal of Econometrics 49, 169–193.

Palm, F., 1977, On univariate time series methods and simultaneous equation econometric models, Journal of Econometrics 5, 379–388.

Quenouille, M.H., 1957, The analysis of multiple time series (Griffin, London).

Sargent, T.J., 1984, Autoregressions, expectations, and advice, American Economic Review 74, 408–415.

Sims, C.A., 1980, Macroeconomics and reality, Econometrica 48, 1–48.

Sims, C.A., 1982, Policy analysis with econometric models, Brookings Papers on Economic Activity, 107–164.

Sims, C.A., 1986, Are forecasting models usable for policy analysis?, Federal Reserve Bank of Minneapolis Quarterly Review 10, Winter, 2–16.

Steel, M.F.J. and J.-F. Richard, 1991, Bayesian multivariate exogeneity analysis, Journal of Econometrics 49, 239–274.

Tiao, G.C. and M.R. Grupe, 1980, Hidden periodic autoregressive–moving average models in time series data, Biometrika 67, 365–373.

Zarnowitz, V. and P.A. Braun, 1990, Major macroeconomic variables and their interrelations, 1882–1983, in: P.A. Klein, ed., Analyzing modern business cycles (M.E. Sharpe, New York, NY) 177–205.

Zellner, A., 1987, Macroeconomics, econometrics and time series analysis, Revista Española de Economia 4, 3–9.

Zellner, A. and F. Palm, 1974, Time series analysis and simultaneous equation econometric models, Journal of Econometrics 2, 17–54.