Lütkepohl, Helmut

Working Paper

Bootstrapping impulse responses in VAR analyses

SFB 373 Discussion Paper, No. 2000,22

Provided in Cooperation with:

Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

Bootstrapping Impulse Responses in VAR Analyses

Helmut Lütkepohl[1]

Humboldt-Universität, Institut für Statistik und Ökonometrie, Spandauer Straße 1, D-10178 Berlin, Germany, email: luetke@wiwi.hu-berlin.de

Abstract. Because the parameters of vector autoregressive processes are often difficult to interpret directly, econometricians use quantities derived from the parameters to disentangle the relationships between the variables. Bootstrap methods are often used for inference on the derived quantities. Alternative bootstrap methods for this purpose are discussed, some related problems are pointed out and proposals are presented to overcome the difficulties at least partly. Some remaining problems are presented.

Keywords. Impulse response, bootstrap, vector autoregression, confidence interval

1   Introduction

In econometric analyses vector autoregressive (VAR) models are established tools for describing the generation process of a set of time series variables. The advantage of these models is that their statistical analysis can be done with standard methods. On the other hand, the direct interpretation of the VAR parameters is often difficult. Therefore, to reveal the interaction between the variables involved, impulse responses and related quantities are used. They are usually determined from the estimated process parameters and are therefore also estimates. Their estimation uncertainty is often visualized by plotting confidence intervals (CIs) together with the impulse response coefficients. These CIs are frequently determined by bootstrap methods. The objective of this lecture is to discuss this practice, present some alternative bootstrap methods and also draw attention to the problems related to the application of these methods in the present context.

Before a more formal treatment of the objects of interest is presented in the next section it may be useful to discuss some characteristics of the variables for which the models are constructed. Economic time series typically have nonzero means, stochastic and/or deterministic trends and seasonal components. Such components have to be allowed for in the models used to describe the data generation process (DGP). For simplicity I will occasionally call variables $I(0)$ if they are generated by a process with stationary stochastic part. If they are not $I(0)$ but become $I(0)$ upon taking first differences, they will be called $I(1)$ variables. In principle $I(0)$ and $I(1)$ variables can have polynomial trends and other deterministic components. Because these components do not affect the quantities of interest in the present context, I will ignore them for simplicity in most of the discussion. Also, some economic variables

cannot be reduced to $I(0)$ variables upon differencing once. Such variables are not considered here again for simplicity, although some of the problems are equally valid or even more so for variables with higher order of integration.

The lecture is organized as follows. I will first present some quantities of interest. Then I will consider alternative bootstrap methods that have been proposed in the recent literature and I will also point out some related problems which are relevant in the present context. Moreover, attempts to overcome the problems will be considered. Most of this discussion is based on systems of $I(0)$ variables generated by purely stochastic stationary processes. Admittedly, this set-up is quite unrealistic. The advantage is, however, that the main problems are most easily presented in this framework. Moreover, it will become apparent that the problems are not mainly due to unusual properties of the variables but arise for the simplest cases as well. In the last section extensions will be discussed.

A word on the notation: The symbol $\mathcal{L}(X)$ will be used to denote the distribution of the random variable or random vector X. The symbol $\xrightarrow{d}$ signifies convergence in distribution. The $(K \times K)$ identity matrix is denoted by I_K and vec is the column stacking operator which stacks the columns of a matrix in a column vector.

2 Quantities of Interest

For a given set of K time series variables $y_t = [y_{1t}, \ldots, y_{Kt}]'$, I consider a VAR process of the form

$$A_0 y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t \tag{2.1}$$

as the basic model in the following. Here the A_i $(i = 0, 1, \ldots, p)$ are $(K \times K)$ coefficient matrices which are restricted in a suitable way so that they can be estimated. The disturbances $u_t = [u_{1t}, \ldots, u_{Kt}]'$ are assumed to be white noise, that is, the u_t are serially independent or at least uncorrelated with zero mean and positive definite covariance matrix Σ_u.

As mentioned in the introduction, in practice there will usually also be other components necessary to describe the DGP of a given set of time series in a suitable manner. For example, there may be deterministic terms such as intercepts, seasonal dummy variables or polynomial trend terms. In addition there may be unmodelled stochastic explanatory variables. For the moment I will not consider those terms because the issues of central importance in the following can be discussed without taking such additional terms into account.

The model (2.1) summarizes the instantaneous and intertemporal relations between the variables. The interactions between the variables are usually difficult to infer directly from the coefficient matrices. Therefore impulse responses and other quantities are often computed from the A_i and Σ_u which help in the interpretation of the model. The following examples are of interest in this context:

- *Forecast error impulse responses*
  Defining

$$\Phi_h = A_0^{-1} \sum_{k=1}^{h} A_k \Phi_{h-k}, \quad h = 1, 2, \ldots, \tag{2.2}$$

  with $\Phi_0 = A_0^{-1}$. The ijth elements of the matrices Φ_s, regarded as a function of s, trace out the expected response of $y_{i,t+s}$ to a unit change in y_{jt}

holding constant all past values of y_t. If $A_0 = I_K$, (2.1) is a reduced form model. In that case the change in y_{it} given $\{y_{t-1}, y_{t-2}, \ldots\}$ is measured by the innovation u_{it} and the elements of Φ_s represent the impulse responses of the components of y_t with respect to the u_t innovations. These impulse responses are sometimes called *forecast error impulse responses* because, for a reduced form model, the u_t are the 1-step ahead forecast errors. Note that generally the $\phi_{ij,h}$ are sums of products of the elements of the A_i. For instance, if $p = 1$ and $A_0 = I_K$ it is easily seen that $\Phi_h = A_1^h$.

— *Orthogonalized impulse responses*
  Denoting by P a lower triangular matrix such that $\Sigma_u = PP'$, the orthogonalized shocks are given by $\varepsilon_t = P^{-1}u_t$ which have unit covariance matrix, that is, $E(\varepsilon_t \varepsilon_t') = I_K$. The orthogonalized impulse responses are the elements of the $\Psi_i = \Phi_i P$ $(i = 0, 1, 2, \ldots)$.

— *Structural impulse responses*
  Since many matrices P exist which satisfy $PP' = \Sigma_u$, considering the orthogonalized impulse responses is to some extent arbitrary. If a priori hypotheses are used to choose a matrix P such that $u_t = P\varepsilon_t$ results in so-called structural impulse responses $\Psi_i = \Phi_i P$ $(i = 0, 1, 2, \ldots)$.

— *Forecast error variance decomposition*
  Let $y_{T+h|T}$ be the optimal h-step forecast in period T. Then the corresponding forecast error can be written as

$$y_{T+h} - y_{T+h|T} = \Psi_0 \varepsilon_{T+h} + \Psi_1 \varepsilon_{T+h-1} + \cdots + \Psi_{h-1} \varepsilon_{T+1}.$$

Denoting the ijth element of Ψ_s by $\psi_{ij,s}$, the kth element of the forecast error vector becomes

$$y_{k,T+h} - y_{k,T+h|T} = \sum_{s=0}^{h-1} (\psi_{k1,s} \varepsilon_{1,T+h-s} + \cdots + \psi_{kK,s} \varepsilon_{K,T+h-s}).$$

Using that the ε_{kt} are contemporaneously and serially uncorrelated and have unit variances by construction, it follows that the corresponding forecast error variance is

$$\sigma_k^2(h) = \sum_{s=0}^{h-1} (\psi_{k1,s}^2 + \cdots + \psi_{kK,s}^2) = \sum_{j=1}^{K} (\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2).$$

The term $(\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2)$ is interpreted as the contribution of variable j to the h-step forecast error variance of variable k. This interpretation makes sense if the ε_{it} can be interpreted as shocks in variable i. Dividing the above terms by $\sigma_k^2(h)$ gives the percentage contribution of variable j to the h-step forecast error variance of variable k.

For detailed discussions of these and further quantities of interest in the present context see Sims (1980, 1981, 1992), Lütkepohl (1990, 1991), Watson (1994), Lütkepohl & Breitung (1997). A crucial property of these quantities from the point of view of our analysis is that they are particular nonlinear functions of the parameters of the model (2.1), say,

$$\phi_{ij,h} = \phi_{ij,h}(A_0, A_1, \ldots, A_p) \quad \text{or} \quad \psi_{ij,h} = \psi_{ij,h}(A_0, A_1, \ldots, A_p, \Sigma_u), \quad (2.3)$$

where $\phi_{ij,h}$ or $\psi_{ij,h}$ represents the response of variable i to an impulse in variable j, h periods ago.

3 Bootstrap Inference

Usually the coefficients of model (2.1) are estimated by some standard procedure such as (pseudo) maximum likelihood (ML), feasible generalized least squares (GLS) or generalized method of moments (GMM). Estimators of the impulse responses are then, for example, obtained as

$$\hat{\phi}_{ij,h} = \phi_{ij,h}(\hat{A}_0, \ldots, \hat{A}_p), \tag{3.1}$$

where the $\hat{A}_0, \ldots, \hat{A}_p$ are the estimated VAR parameter matrices. The properties of the estimator $\hat{\phi}_{ij,h}$ then follow from those of $\hat{A}_0, \ldots, \hat{A}_p$. To assess the sampling uncertainty of the quantities of interest, confidence intervals (CIs) are sometimes determined on the basis of the asymptotic distributions. Alternatively, bootstrap methods are often used for this purpose, because these methods occasionally lead to more reliable small sample inference than CIs based on standard asymptotic theory. It may be worth noting, however, that the bootstrap is also justified by asymptotic arguments only. In the following, alternative bootstrap approaches will be considered for setting up CIs for impulse responses. They are all residual based bootstraps using the following general algorithm which assumes a sample of size T plus some presample values.

(1) Estimate the parameters of the model in (2.1) and denote the estimation residuals by $\hat{u}_t$.

(2) Compute centered residuals $\hat{u}_1 - \bar{u}_., \ldots, \hat{u}_T - \bar{u}_.$, where $\bar{u}_. = T^{-1} \sum \hat{u}_t$, and generate bootstrap residuals $u_1^*, \ldots, u_T^*$ by randomly drawing with replacement from the centered residuals.

(3) Compute bootstrap time series recursively as

$$y_t^* = \hat{A}_0^{-1} \left(\hat{A}_1 y_{t-1}^* + \cdots + \hat{A}_p y_{t-p}^* + u_t^* \right), \quad t = 1, \ldots, T,$$

where $(y_{-p+1}^*, \ldots, y_0^*) = (y_{-p+1}, \ldots, y_0)$.

(4) Reestimate the parameters $A_0, \ldots, A_p$ based on the bootstrap time series.

(5) Based on the parameter estimates obtained in Stage (4), calculate a bootstrap version of the statistic of interest, say $\hat{\phi}^*$.

(6) Repeat Steps (2) - (5) a large number of times (say N times).

This algorithm is based on an assumed VAR(p) model. If there is uncertainty with respect to specific aspects of the model like, for instance, the VAR order, a block bootstrap may be used which is based on the original observations rather than the model residuals (see, e.g., Li & Maddala (1996) for details). Because construction of a suitable model for the DGP is often the main purpose of an analysis, using the block bootstrap is not an obvious procedure in the present context, however, and will not be discussed further.

In the following I use the symbols ϕ, $\hat{\phi}$ and $\hat{\phi}^*$ to denote some general impulse response coefficient, its estimator implied by the estimators of the model coefficients and the corresponding bootstrap estimator, respectively. The following bootstrap CIs have, for instance, been considered in the literature:

– *Standard percentile interval*
The most common method in setting up CIs for impulse responses in practice is to use the interval

$$CI_S = \left[s^*_{\gamma/2}, s^*_{(1-\gamma/2)} \right],$$

where $s^*_{\gamma/2}$ and $s^*_{(1-\gamma/2)}$ are the $\gamma/2$- and $(1-\gamma/2)$-quantiles, respectively, of the bootstrap distribution $\mathcal{L}(\hat{\phi}^* \mid y_{-p+1}, \ldots, y_0, \ldots, y_T)$. The interval CI_S is the percentile confidence interval described, e.g., by Efron & Tibshirani (1993).

– *Hall's percentile interval*
Hall (1992) uses the usual bootstrap analogy,

$$\mathcal{L}(\hat{\phi} - \phi) \approx \mathcal{L}(\hat{\phi}^* - \hat{\phi} \mid y_{-p+1}, \ldots, y_0, \ldots, y_T),$$

to derive the interval

$$CI_H = \left[\hat{\phi} - t^*_{(1-\gamma/2)}, \hat{\phi} - t^*_{\gamma/2} \right],$$

where $t^*_{\gamma/2}$ and $t^*_{(1-\gamma/2)}$ are the $\gamma/2$- and $(1-\gamma/2)$-quantiles of $\mathcal{L}(\hat{\phi}^* - \hat{\phi} \mid y_{-p+1}, \ldots, y_0, \ldots, y_T)$, respectively.

– *Hall's studentized interval*
In some situations, using a studentized statistic $(\hat{\phi} - \phi)/(\widehat{\text{var}}(\hat{\phi}))^{1/2}$ for constructing confidence intervals may be advantageous. In that case bootstrap quantiles from the distribution of $(\hat{\phi}^* - \hat{\phi})/(\widehat{\text{var}}(\hat{\phi}^*))^{1/2}$ are used to construct an interval

$$CI_{SH} = \left[\hat{\phi} - t^{**}_{(1-\gamma/2)}(\widehat{\text{var}}(\hat{\phi}))^{1/2}, \ \hat{\phi} - t^{**}_{\gamma/2}(\widehat{\text{var}}(\hat{\phi}))^{1/2} \right].$$

In this approach the variances are estimated by a bootstrap within each bootstrap replication.

A number of refinements and modifications of these intervals exist. Some of them will be mentioned in the next section where important properties of the present intervals are discussed.

4 Properties and Problems

A desirable property of a CI is that it has the nominal confidence content at least asymptotically. The bootstrap CIs mentioned in the previous section have this property under certain conditions. Roughly speaking, if $\mathcal{L}(\sqrt{T}(\hat{\phi} - \phi))$ converges as $T \to \infty$, $\mathcal{L}(\sqrt{T}(\hat{\phi}^* - \hat{\phi}) \mid y_{-p+1}, \ldots, y_0, \ldots, y_T)$ converges to the same limit distribution under suitable conditions (e.g., Hall (1992)). It follows immediately that in this case CI_H has the correct size asymptotically, that is, $Pr\,(\phi \in CI_H) \to 1 - \gamma$ as $T \to \infty$ and, hence, Hall's percentile method is asymptotically correct. Similar arguments can be used to establish the asymptotic validity of the CI_{SH} interval. The same can also be shown for CI_S if the limiting distribution corresponding to $\mathcal{L}(\sqrt{T}(\hat{\phi} - \phi))$ is symmetric about zero, for instance, if it is zero mean normal. Roughly

speaking, CI_S works with an implicit unbiasedness assumption for $\hat{\phi}$. However, if the distribution of $\hat{\phi}$ is not centered at ϕ, CI_S will generally not have the desired confidence content even asymptotically (see also Efron & Tibshirani (1993) and Benkwitz, Lütkepohl & Neumann (2000) for a more detailed discussion of this point). Hence, we have to consider the distribution of the quantities of interest.

Assuming that the $\hat{A}_i$ have an asymptotic normal distribution,

$$\sqrt{T}\text{vec}([\hat{A}_0, \ldots, \hat{A}_p] - [A_0, \ldots, A_p]) \xrightarrow{d} N(0, \Sigma_{\hat{A}}), \qquad (4.1)$$

using the delta method (e.g., Serfling (1980)), $\hat{\phi}$ can be shown to have an asymptotic normal distribution as well,

$$\sqrt{T}(\hat{\phi} - \phi) \xrightarrow{d} N(0, \sigma_{\hat{\phi}}^2), \qquad (4.2)$$

where

$$\sigma_{\hat{\phi}}^2 = \frac{\partial \phi}{\partial \alpha'} \Sigma_{\hat{A}} \frac{\partial \phi}{\partial \alpha} \qquad (4.3)$$

with $\alpha = \text{vec}[A_0, \ldots, A_p]$, and $\partial \phi / \partial \alpha$ denotes a vector of first order partial derivatives. The result (4.2) holds if $\sigma_{\hat{\phi}}^2$ is nonzero which is a crucial condition for asymptotic inference to work in the usual way. Note that $\Sigma_{\hat{A}}$ may be singular if there are constraints on the coefficients or if the variables are integrated and/or cointegrated (see Lütkepohl (1991, Chapter 11)). However, even if $\Sigma_{\hat{A}}$ is nonsingular, $\sigma_{\hat{\phi}}^2$ may be zero because the partial derivatives in (4.3) may be zero. In fact, they will usually be zero in parts of the parameter space because an impulse response coefficient ϕ generally consists of sums of products of the VAR parameters and, hence, the partial derivatives will also be sums of products of such coefficients which may be zero, of course.

To see this consider a simple, stable, univariate AR(1) process $y_t = \alpha y_{t-1} + u_t$ with $|\alpha| < 1$. In that case, forecast error impulse responses are of the form $\phi_h = \alpha^h$. If $\hat{\alpha}$ is the usual LS estimator, it is well-known that $\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{d} N(0, 1 - \alpha^2)$ under general conditions (Anderson (1959)) and, hence, $\sqrt{T}(\hat{\phi}_h - \phi_h) \xrightarrow{d} N(0, h^2 \alpha^{2(h-1)}(1 - \alpha^2))$ for $\alpha \neq 0$. If $\alpha = 0$, however, $\sqrt{T}(\hat{\phi}_h - \phi_h) = \sqrt{T}\hat{\alpha}^h \xrightarrow{d} 0$ for $h > 1$. In other words, $(\hat{\phi}_h - \phi_h)$ converges to zero more rapidly than $T^{-1/2}$ so that the convergence rate is not constant for all points in the parameter space. This problem is aggravated for integrated processes with $\alpha = 1$ and for higher order or higher dimensional processes.

In fact, in the simple AR(1) example it is easy to see that CI_S has zero coverage probability for $\alpha = 0$ and even h, because in that case, denoting by $\hat{\alpha}^*$ the estimator of α based on the bootstrap time series, we have that $(\hat{\alpha}^*)^h$ will be strictly positive with probability 1. Consequently, using simply the relevant lower and upper percentage points of the bootstrap distribution results in an interval which cannot include zero. Hence, $\alpha^h = \phi_h = 0$ is not included. Although the other bootstrap CIs do not have this problem, they may not have the correct coverage probability asymptotically (see Benkwitz, Lütkepohl & Neumann (2000) for details).

The latter authors have also performed a Monte Carlo experiment to investigate the performance of bootstrap CIs in finite samples. Some of their

Table 4.1. Relative coverage frequencies of nominal 95% CIs for impulse responses ϕ_h

	$\alpha = 0$	$\alpha = .20$	$\alpha = .50$	$\alpha = .90$	$\alpha = .99$
			$h = 1$		
CI_S	.954	.953	.953	.891	.776
CI_H	.933	.929	.937	.890	.929
CI_{SH}	.942	.942	.945	.915	.928
			$h = 2$		
CI_S	.000	.982	.953	.891	.776
CI_H	.982	.703	.876	.882	.925
CI_{SH}	.979	.788	.956	.920	.933
			$h = 3$		
CI_S	.954	.953	.953	.891	.776
CI_H	1.000	.676	.821	.866	.919
CI_{SH}	1.000	.764	.958	.924	.935
			$h = 4$		
CI_S	.000	.982	.953	.891	.776
CI_H	.976	.620	.786	.855	.905
CI_{SH}	.979	.764	.962	.929	.937

(CI_{SH} figures extracted from Tables III - VI of Benkwitz, Lütkepohl & Neumann (2000). Other figures simulated based on the same generated time series.)

simulation results are replicated in Table 4.1 together with own results. The figures are based on the AR(1) example, a sample size $T = 100$ and $N = 2000$ bootstrap replications in each simulation run. The relative coverage frequencies shown in Table 4.1 are obtained for 1000 replications of the experiment.

Obviously, CI_S fails completely for $\alpha = 0$ and $h = 2$ and 4, that is, the true ϕ_h never falls in CI_S in these cases. For odd h or $\alpha > 0$, the coverage of CI_S improves considerably. It deteriorates, however, if α approaches 1 and, hence, the instability region. In contrast, CI_H and CI_{SH} perform well in terms of coverage for $\alpha = 0$ and $h = 1, 2, 3, 4$. They sometimes even have a larger coverage probability than desired. They do not perform very well in some of the other cases, however. For example for $h = 4$ and $\alpha = .20$, the actual coverage frequencies of around 60% for CI_H and of 76.4% of CI_{SH} are not satisfactory, although it is apparent that CI_{SH} performs better than CI_H in general. For α close to 1, both CI_H and CI_{SH} have a better coverage than CI_S.

In summary, the three types of CIs considered here suffer from both asymptotic as well as small sample deficiencies. In particular, there are points in the parameter space where the asymptotic theory fails and, thus, it cannot

be used to justify the bootstrap. Moreover, some intervals perform poorly in small samples even for asymptotically unproblematic true parameter values. For cases where the coverage is correct asymptotically, small sample adjustments have been proposed in the literature. For instance, Hall (1992) argues that an iterative bootstrap procedure may result in more accurate CIs. In that method a further layer of bootstrap samples is drawn from each original bootstrap sample. Then CIs are computed from each of the second stage bootstrap samples and these CIs are used to estimate the actual coverage by checking how often the original estimate falls within these intervals. Then the original CI is furnished with a correction factor to adjust the coverage. This procedure can be used iteratively by repeating it more than once.

Small sample distortions of CI_S may also reflect biased estimation procedures. In those cases a bias correction may be used. A possible procedure for the presently considered cases was proposed and investigated by Kilian (1998a). Although these modifications help to improve the coverage probabilities in some small sample situations which are such that the standard asymptotic theory works, they do not overcome the problems for those cases where the usual $\sqrt{T}$-asymptotics break down.

5 Possible Solutions

There are at least three possible strategies to overcome the problems resulting from the different rates of convergence in the parameter space. First, one may consider bootstrap procedures that adjust to the kind of singularity in the asymptotic distribution which we have to deal with here. A couple of different proposals of this sort are discussed in Benkwitz, Lütkepohl & Neumann (2000). For instance, subsampling may be used to estimate the convergence rate in addition to the model parameters. These and other methods were shown to have drawbacks, however, in empirical applications. Either they are not very practical for processes of realistic dimension and autoregressive order or they do not perform well in samples of typical size.

A second possibility to tackle the singularity problem is to single out and eliminate the points where problems occur before an impulse response analysis is carried out. In the present case this proposal amounts to determining all zero coefficients in a first stage of the analysis and enforcing the resulting zero restrictions in the next stage where the model underlying the impulse response analysis is estimated and used for computing impulse responses. This solution is, for instance, considered by Benkwitz, Lütkepohl & Wolters (2000). Problematic in this approach may be the uncertainty with respect to the actual zero restrictions.

A third way out of the singularity problem is to consider a different type of modelling approach based on the assumption of a potentially infinite VAR order. So far it has been assumed that a given finite order VAR model is considered. In practice, it is usually unknown a priori which model provides a suitable representation of the DGP. Therefore the model order and other restrictions are chosen by some kind of statistical procedure. Kilian (1998b) proposes to use a bootstrap procedure at this stage as well. In any case, such a preliminary analysis introduces additional uncertainty into the analysis. An alternative approach is therefore to explicitly allow the original model to be an infinite order VAR process and to integrate the model choice directly into the estimation procedure by assuming that the model order is increased when more sample information becomes available. In other word, the model order is assumed to approach infinity with the sample size. A suitable asymptotic

theory is developed by Lütkepohl (1988, 1996), Lütkepohl & Poskitt (1991, 1996), Lütkepohl & Saikkonen (1997) and Saikkonen & Lütkepohl (1995, 1996) based on work by Lewis & Reinsel (1985) and Saikkonen (1992). It turns out that this asymptotic theory avoids the kind of singularity in the asymptotic distribution which causes the failure of the bootstrap procedures. On the other hand, the greater generality of the model results in an inefficiency relative to the model with finite fixed order.

6   Extensions and Open Problems

As mentioned earlier, in practice deterministic terms are often used in VAR processes. Such terms can be added without affecting the previous results because they do not enter the computation of the impulse response coefficients, that is, impulse responses are computed for the stochastic part only anyway. For $I(0)$ processes, the estimates of the deterministic part will be asymptotically independent of the stochastic part if the model is specified in a suitable way. Hence, from the point of view of asymptotic theory no new problems arise. Of course, in small samples the inclusion of deterministic terms can matter to some extent. Moreover, additional stochastic exogenous variables and their impact on the endogenous variables may be of interest. In that case, dynamic multipliers may be considered which are closely related to impulse responses. In principle, the analysis of the former quantities can proceed in a similar way as for impulse responses.

A different situation arises if $I(1)$ variables are also included in the system under consideration. In that case there may be singularities in the limiting distribution of the VAR parameters due to unit roots. Therefore, the singularity problem is aggravated. Unfortunately, in that case it cannot be overcome simply by imposing zero restrictions on the parameters because it arises from the superefficient estimation of cointegration parameters or, alternatively, of certain linear combinations of VAR parameters. One possible way to get around the resulting problems for bootstrapping impulse responses may be a detailed modelling of the cointegration relations and taking into account the implied singularities in the asymptotic distribution. In fact, in that case it may be possible to estimate the cointegration relations in a first round and fix the associated parameters in the bootstrap procedure. Such an approach is discussed by Benkwitz, Lütkepohl & Wolters (2000) who also present detailed examples. Unfortunately, a complete asymptotic justification of such a procedure seems to be missing to date.

Sims & Zha (1999) point out that the individual impulse responses are not independent. Consequently, reporting CIs for individual impulse response coefficients may give a misleading impression of the actual uncertainty in the estimated impulse response function. In other words, the band obtained by drawing individual $(1 - \gamma)100\%$ CIs around each coefficient separately is not a confidence band with confidence content $(1 - \gamma)100\%$. Sims & Zha propose a method that can help in getting a better picture of the overall uncertainty in the estimated impulse response functions.

So far I have discussed linear models only. In these models impulse responses are especially easy to analyze because the response to an innovation does not depend on the state of the system. Moreover, the direction or the size of a shock do not have an impact on the shape of the response. Therefore it is easy to represent the essential features of the dynamic interactions between the variables in a relatively small set of impulse response functions. On the other hand, it is clear that these invariance properties are not realistic in all situations of practical interest. As a consequence, nonlinear dynamic

models and corresponding extensions of impulse response functions have been proposed in the literature (e.g., Gallant, Rossi & Tauchen (1993), Pesaran & Shin (1998)). In general the inference problems observed in the simpler linear models will be aggravated if more complicated models are considered.

References

Anderson, T. W. (1959). On asymptotic distribution of estimates of parameters of stochastic difference equations. *The Annals of Mathematical Statistics*, 30, 676–687.

Benkwitz, A., H. Lütkepohl & M. H. Neumann (2000). Problems related to confidence intervals for impulse responses of autoregressive processes. *Econometric Reviews*, 19, 69–103.

Benkwitz, A., H. Lütkepohl & J. Wolters (2000). Comparison of bootstrap confidence intervals for impulse responses of German monetary systems. *Macroeconomic Dynamics*, forthcoming.

Efron, B. & R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Gallant, A.R., P.E. Rossi & G. Tauchen (1993). Nonlinear dynamic structures. *Econometrica*, 61, 871–907.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Kilian, L. (1998a). Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics*, 80, 218–230.

Kilian, L. (1998b). Accounting for lag order uncertainty in autoregressions: The endogenous lag order bootstrap. *Journal of Time Series Analysis*, 19, 531–548.

Lewis, R. & G. C. Reinsel (1985). Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis*, 16, 393–411.

Li, H. & G.S. Maddala (1996). Bootstrapping time series models. *Econometric Reviews*, 15, 115–158.

Lütkepohl, H. (1988). Asymptotic distribution of the moving average coefficients of an estimated vector autoregressive process. *Econometric Theory*, 4, 77–85.

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.

Lütkepohl, H. (1991). *Introduction to Multiple Times Series Analysis*. Berlin: Springer–Verlag.

Lütkepohl, H. (1996). Testing for nonzero impulse responses in vector autoregressive processes. *Journal of Statistical Planning and Inference*, 50, 1–20.

Lütkepohl, H. & J. Breitung (1997). Impulse response analysis of vector autoregressive processes, in: C. Heij, J.M. Schumacher, B. Hanzon and C. Praagman (eds.), *System Dynamics in Economic and Financial Models*, Chichester: John Wiley, 299–326.

Lütkepohl, H. & D.S. Poskitt (1991). Estimating orthogonal impulse responses via vector autoregressive models. *Econometric Theory*, 7, 487–496.

Lütkepohl, H. & D.S. Poskitt (1996). Testing for causation using infinite order vector autoregressive processes. *Econometric Theory*, 12, 61–87.

Lütkepohl, H. & P. Saikkonen (1997). Impulse response analysis in infinite order cointegrated vector autoregressive processes. *Journal of Econometrics*, 81, 127–157.

Pesaran, H.H. & Y. Shin (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58, 17–29.

Saikkonen, P. (1992). Estimating and testing of cointegrated systems by an autoregressive approximation. *Econometric Theory*, 8, 1–27.

Saikkonen, P. & H. Lütkepohl (1995). Asymptotic inference on nonlinear functions of the coefficients of infinite order cointegrated VAR processes. Discussion Paper 66, SFB 373, Humboldt University Berlin.

Saikkonen, P. & H. Lütkepohl (1996). Infinite order cointegrated vector autoregressive processes: Estimation and inference. *Econometric Theory*, 12, 814–844.

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

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

Sims, C. A. (1981). An autoregressive index model for the U.S. 1948-1975, in: J. Kmenta & J.B. Ramsey (eds.), *Large-Scale Macro-Econometric Models*, Amsterdam: North-Holland, 283-327.

Sims, C.A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, 36, 975–1000.

Sims, C.A. & T. Zha (1999). Error bands for impulse responses. *Econometrica*, 67, 1113–1155.

Watson, M. W. (1994). Vector autoregressions and cointegration, in: Engle, R.F. and D.L. McFadden (eds.), *Handbook of Econometrics*, Vol. IV, New York: Elsevier.