Macroeconomic Dynamics, **5**, 2001, 81–100. Printed in the United States of America.

COMPARISON OF BOOTSTRAP CONFIDENCE INTERVALS FOR IMPULSE RESPONSES OF GERMAN MONETARY SYSTEMS

ALEXANDER BENKWITZ AND HELMUT LÜTKEPOHL
Humboldt-Universität Zu Berlin, Institut für Statistik und Ökonometrie

JÜRGEN WOLTERS
Freie Universität Berlin, Institut für Statistik und Ökonometrie

It is argued that standard impulse response analysis based on vector autoregressive models has a number of shortcomings. Although the impulse responses are estimated quantities, measures for sampling variability such as confidence intervals sometimes are not provided. If confidence intervals are given, they often are based on bootstrap methods with dubious theoretical properties. These problems are illustrated using two German monetary systems. Proposals are made for improving current practice. Special emphasis is placed on systems with cointegrated variables.

Keywords: Impulse Response, Bootstrap Methods, Money Demand System, Monetary Policy

1. INTRODUCTION

Impulse responses are standard tools in vector autoregressive (VAR) analyses. In this context an economic system of interest is described by a VAR model that is estimated from the available time-series data in unrestricted form or with various types of structural and statistical restrictions imposed. There are a number of problems related to commonly applied procedures. First, impulse responses are computed from estimated coefficients and are therefore also estimates. Although this fact is recognized in the literature, sometimes only the point estimates are plotted and the relation of the variables involved is interpreted on the basis of these point estimates without properly taking into account the estimation variability [e.g., Sims (1992), Hendry and Mizon (1998), Pesaran and Shin (1998)]. In another part of the literature, the estimation uncertainty of impulse responses is assessed by setting up confidence intervals (CI's). In many studies, however, it

We thank Michael H. Neumann and two anonymous referees for helpful comments. Moreover, we are grateful to the Deutsche Forschungsgemeinschaft, SFB 373, and the European Commission under the Training and Mobility of Researchers Programme (Contract No. ERBFMRXCT980213) for financial support. Address correspondence to: Prof. H. Lütkepohl, Humboldt-Universität, Institut für Statistik und Ökonometrie, Spandauer Straße 1, D-10178 Berlin, Germany; e-mail: luetke@wiwi.hu-berlin.de.

 1365-1005/01 $9.50

was found that the CI's are rather wide and, hence, the impulse responses are not very informative, so that nothing much can be said about the actual underlying relations. Clearly, this reflects the substantial sampling variability in the estimated VAR parameters, which in turn is a consequence of estimating these quantities in a largely unrestricted model with many parameters.

Another potential problem in this context is that the CI's for the impulse responses often are based on bootstrap methods. It has been argued by Benkwitz, Lütkepohl, and Neumann (2000) (henceforth BLN) that the usual bootstrap procedure used in this context can fail completely by producing CI's with actual coverage probability of zero, regardless of the desired nominal confidence level. In other words, the bootstrap CI's may give a grossly distorted impression of the range of likely impulse responses for a VAR model.

The purpose of this paper is to illustrate and discuss the importance of these problems for applied work. We use two small German monetary systems and show that it is crucial to take into account the estimation uncertainty when interpreting impulse responses in the context of dynamic econometric models. We compare the commonly used bootstrap methods for determining CI's of impulse responses to other methods that currently are not very popular in the macroeconometric literature but have been proposed in the bootstrap literature. We point out that a method proposed by Hall (1992) is advantageous in some respects. Moreover, we show that imposing restrictions on the short-term dynamics of a system can reduce the length of the CI's substantially, which in turn can lead to a more informative picture of the dynamic interactions between the variables of the system under consideration. In our analysis, we focus on vector error correction models (VECM's) and we will pay special attention to the treatment of cointegration relations.

The paper has the following structure. The general framework of the analysis is presented in the next section and inference on impulse responses is considered in Section 3. In particular, alternative methods for computing bootstrap CI's for impulse responses are discussed. These methods are applied and compared within two small monetary systems for Germany in Section 4. Conclusions are drawn in Section 5.

The following notation is used throughout: $\mathcal{L}(X)$ denotes the distribution of the random variable X. The natural logarithm is abbreviated as ln and Δ is the differencing operator defined such that for a time-series variable y_t, $\Delta y_t = y_t - y_{t-1}$. Nonstationary variables that become stationary upon differencing once are referred to as $I(1)$ variables. The indicator function is denoted by $\mathcal{I}(\cdot)$.

2. ANALYSIS OF VAR PROCESSES

Many macroeconomic analyses are based on linear dynamic models of the type

$$A_0 \boldsymbol{y}_t = A_1 \boldsymbol{y}_{t-1} + \cdots + A_p \boldsymbol{y}_{t-p} + \Psi \boldsymbol{x}_t + \Xi D_t + \boldsymbol{u}_t, \tag{1}$$

where $\boldsymbol{y}_t = [y_{1t}, \ldots, y_{Kt}]'$ is a K-dimensional vector of observable endogenous variables; the A_i $(i = 0, 1, \ldots, p)$ are $(K \times K)$ coefficient matrices; $\boldsymbol{x}_t$ represents

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

a vector of N unmodeled observable variables; D_t contains all deterministic terms such as seasonal dummy variables, intercept, and polynomial trend terms; Ψ and Ξ are also coefficient matrices; and $\boldsymbol{u}_t = [u_{1t}, \ldots, u_{Kt}]'$ is a white-noise process, that is, the $\boldsymbol{u}_t$ are serially uncorrelated or independent with zero mean and nonsingular (positive definite) covariance matrix Σ_u. The model (1) is somewhat more general than the typical pure VAR model in that it may contain unmodeled variables whereas in standard VAR analyses all stochastic variables are treated as endogenous. We still refer to (1) as our basic VAR model. The maximum lag length p of the endogenous variables usually is referred to as the order of the VAR process and the process is briefly called a VAR(p). The process may be stationary or it may contain $I(1)$ variables and r cointegrating relations, where $0 < r < K$. In the latter case, it is often written as a VECM,

$$\Gamma_0 \Delta \boldsymbol{y}_t = \alpha\beta' \boldsymbol{y}_{t-1} + \Gamma_1 \Delta \boldsymbol{y}_{t-1} + \cdots + \Gamma_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \Psi \boldsymbol{x}_t + \Xi D_t + \boldsymbol{u}_t, \quad \textbf{(2)}$$

where the Γ_j $(j = 0, 1, \ldots, p-1)$ are the short-run parameter matrices, α is the $(K \times r)$ loading matrix and β is a $(K \times r)$ matrix containing r linearly independent cointegration relations. In the examples in Section 4 the exogenous variables $\boldsymbol{x}_t$ are stationary variables.

Regardless of the stationarity properties, the model in (1) or (2) summarizes the instantaneous and intertemporal relations between the variables. The exact form of these relations is usually difficult to see directly from the coefficients, especially if there are only just identifying restrictions on the short-term parameters Γ_i $(i = 0, 1, \ldots, p-1)$. Therefore, impulse response functions often are computed that represent the marginal responses of the endogenous variables of the system to an impulse in one of the endogenous variables. These may be regarded as conditional forecasts of the endogenous variables given that they have been zero up to time 0 when an impulse in one of the variables occurs. Depending on the kind of impulse hitting the system, there are various different impulse responses that have been used for interpreting VAR models. For detailed discussions, see Sims (1980, 1981), Lütkepohl (1990, 1991), Watson (1994), and Lütkepohl and Breitung (1997). The important 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 in (1) or (2), for example,

$$\phi_{ij,h} = \phi_{ij,h}(A_0, A_1, \ldots, A_p) = \phi_{ij,h}(\alpha, \beta, \Gamma_0, \Gamma_1, \ldots, \Gamma_{p-1}), \qquad \textbf{(3)}$$

where $\phi_{ij,h}$ represents the response of variable i to an impulse in variable j, h periods ago. Precise formulas for different versions of impulse responses may be found in Lütkepohl (1991, Ch. 2) or Lütkepohl and Breitung (1997), for instance. Because the VECM in (2) can always be written in the equivalent levels form in (1) and vice versa and because our example models in Section 4 are VECM's we focus on the latter version in the following in order to minimize repetition. The VECM is also the more convenient model form for discussing the treatment of cointegration relations.

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

3. INFERENCE ON IMPULSE RESPONSES

Usually the coefficients of the model in (2) are estimated by some standard procedure such as (pseudo) maximum likelihood (ML) or feasible generalized least squares (GLS), possibly estimating the cointegration parameters in a first stage and keeping them fixed in estimating the other parameters. Estimators of the impulse responses then are obtained as

$$\hat{\phi}_{ij,h} = \phi_{ij,h}(\hat{\alpha}, \hat{\beta}, \hat{\Gamma}_0, \hat{\Gamma}_1, \ldots, \hat{\Gamma}_{p-1}), \tag{4}$$

where the $\hat{\alpha}, \hat{\beta}, \hat{\Gamma}_0, \hat{\Gamma}_1, \ldots, \hat{\Gamma}_{p-1}$ are the estimated VECM parameter matrices. Under general assumptions, the resulting impulse responses have asymptotic normal distributions which may be used for constructing CI's. In practice, bootstrap methods often are used for this purpose because these methods occasionally lead to more reliable small sample inference than CI's based on standard asymptotic theory. However, we want to emphasize that both approaches, standard asymptotics and the bootstrap, are based on asymptotic arguments.

The analytical expressions of the asymptotic variances of the impulse response coefficients are rather complicated. Using the bootstrap for setting up CI's, the precise expressions of the variances are not needed and, hence, deriving the analytical expressions can be avoided. In the following, we discuss some methods that have been proposed in this context.

The following bootstrap method is considered:

1. Estimate the parameters of the model in (2) by a suitable procedure.
2. Generate bootstrap residuals $\boldsymbol{u}_1^*, \ldots, \boldsymbol{u}_T^*$ by randomly drawing with replacement from the set of estimated and recentered residuals, $\{\hat{\boldsymbol{u}}_1 - \bar{\boldsymbol{u}}_., \ldots, \hat{\boldsymbol{u}}_T - \bar{\boldsymbol{u}}_.\}$, where $\hat{\boldsymbol{u}}_t = \hat{\Gamma}_0 \Delta \boldsymbol{y}_t - \hat{\alpha}\hat{\beta}' \boldsymbol{y}_{t-1} - \hat{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} - \cdots - \hat{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} - \hat{\Psi} \boldsymbol{x}_t - \hat{\Xi} D_t$, and $\bar{\boldsymbol{u}}_. = T^{-1} \sum \hat{\boldsymbol{u}}_t$.
3. Set $(\boldsymbol{y}_{-p+1}^*, \ldots, \boldsymbol{y}_0^*) = (\boldsymbol{y}_{-p+1}, \ldots, \boldsymbol{y}_0)$ and construct bootstrap time series recursively using the levels representation given in (1),
$$\boldsymbol{y}_t^* = \hat{A}_0^{-1}\left(\hat{A}_1 \boldsymbol{y}_{t-1}^* + \cdots + \hat{A}_p \boldsymbol{y}_{t-p}^* + \hat{\Psi} \boldsymbol{x}_t + \hat{\Xi} D_t + \boldsymbol{u}_t^*\right), \qquad t = 1, \ldots, T.$$
4. Reestimate the parameters $\Gamma_0, \Gamma_1, \ldots \Gamma_{p-1}, \Psi, \Xi, \alpha, \beta$ from the generated data.
5. Calculate a bootstrap version of the statistic of interest, for example, $\hat{\phi}_{ij,h}^*$, based on the parameter estimates obtained in Stage 4.

In Stage 4, where the bootstrap estimates are computed, there are two alternative ways to do so. The first possibility is to use the same estimation method in each bootstrap replication that was used in estimating the VECM coefficients from the original data. In this procedure the cointegration matrix β is reestimated for each bootstrap sample. Alternatively, one may argue that the β matrix is estimated superconsistently from the original data and therefore is treated as known and fixed in the bootstrap replications. We explore these two possibilities in the context of the examples in Section 4.

In the following, we use the symbols ϕ, $\hat{\phi}_T$, and $\hat{\phi}_T^*$ to denote some general impulse response coefficient, its estimator implied by the estimators of the model coefficients, and the corresponding bootstrap estimator, respectively. The subscript T indicates the sample size.

The most commonly used method in setting up CI's for impulse responses in practice proceeds by using $\gamma/2$- and $(1-\gamma/2)$-quantiles, for example, $s^*_{\gamma/2}$ and $s^*_{(1-\gamma/2)}$, respectively, of the bootstrap distribution $\mathcal{L}(\hat{\phi}^*_T|\boldsymbol{y}_{-p+1},\dots,\boldsymbol{y}_0,\dots,\boldsymbol{y}_T;\boldsymbol{x}_1,\dots,\boldsymbol{x}_T)$, and defining

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

The interval CI_S is the percentile confidence interval described, for example, by Efron and Tibshirani (1993). Those authors point out, however, that it may not have the desired coverage probability. This problem occurs, for example, if $\hat{\phi}_T$ is a biased estimator of ϕ. In that case, the bootstrap distribution may be asymptotically centered at ϕ plus a bias term and, hence, CI_S is a $(1-\gamma)100\%$ CI for the latter quantity and may have a grossly distorted level as a CI for ϕ. To fix this drawback, modifications of CI_S have been proposed in the literature. In the context of impulse response analysis, Kilian (1998) has suggested a method to reduce the problem. In the bootstrap literature [see, e.g., Hall (1992)], other modifications have been proposed, some of which are presented in the following.

Let $t^*_{\gamma/2}$ and $t^*_{(1-\gamma/2)}$ be the $\gamma/2$- and $(1-\gamma/2)$-quantiles of

$$\mathcal{L}\left(\hat{\phi}^*_T - \hat{\phi}_T \,\middle|\, \boldsymbol{y}_{-p+1},\dots,\boldsymbol{y}_0,\dots,\boldsymbol{y}_T;\boldsymbol{x}_1,\dots,\boldsymbol{x}_T\right),$$

respectively. According to the usual bootstrap analogy,

$$\mathcal{L}(\hat{\phi}_T - \phi) \approx \mathcal{L}\left(\hat{\phi}^*_T - \hat{\phi}_T \,\middle|\, \boldsymbol{y}_{-p+1},\dots,\boldsymbol{y}_0,\dots,\boldsymbol{y}_T;\boldsymbol{x}_1,\dots,\boldsymbol{x}_T\right),$$

one gets the interval

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

Hall (1992) calls this CI the "percentile interval." Therefore, in the following we refer to the method leading to CI_H as Hall's percentile method, whereas the method underlying CI_S is referred to as the standard method. If $\mathcal{L}[\sqrt{T}(\hat{\phi}^*_T - \hat{\phi}_T)\,|\,\boldsymbol{y}_{-p+1},\dots,\boldsymbol{y}_0,\dots,\boldsymbol{y}_T;\boldsymbol{x}_1,\dots,\boldsymbol{x}_T]$ has the same limit distribution as $\mathcal{L}[\sqrt{T}\,(\hat{\phi}_T - \phi)]$, then it follows immediately that 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.

It is well established in the bootstrap literature that the quality of the bootstrap approximation of the distribution of a general statistic $\hat{\mu}_T$, for example, can be improved by reducing its dependence on the unknown distribution that governs the data generating process. For example, with respect to the sample mean of i.i.d. random variables it is well known that studentizing leads to a better rate of approximation by the bootstrap [see, e.g., Hall (1992)]. Therefore, it may be advantageous to use a studentized statistic $(\hat{\phi}_T - \phi)/\sqrt{\widehat{\text{var}}(\hat{\phi}_T)}$ as a basis for constructing confidence intervals. Hence in the present context it may be useful

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

to determine a bootstrap quantile based on the statistic $(\hat{\phi}_T^* - \hat{\phi}_T)/\sqrt{\widehat{\text{var}}(\hat{\phi}_T^*)}$. In this approach, the variances also are estimated by a bootstrap; that is,

$$\widehat{\text{var}}(\hat{\phi}_T) = \frac{1}{B^* - 1} \sum_{i=1}^{B^*} \left(\hat{\phi}_T^{*,i} - \overline{\hat{\phi}_T^*}\right)^2$$

and

$$\widehat{\text{var}}\left(\hat{\phi}_T^*\right) = \frac{1}{B^{**} - 1} \sum_{i=1}^{B^{**}} \left(\hat{\phi}_T^{**,i} - \overline{\hat{\phi}_T^{**}}\right)^2,$$

where $\hat{\phi}_T^{**,i}$ is obtained by a double bootstrap. That is, pseudo-data are generated according to a process obtained on the basis of the bootstrap systems parameters and B^* and B^{**} are the respective numbers of bootstrap replications in the first and second stages [see Hall (1992) for details].

Let $t_{\gamma/2}^{**}$ and $t_{(1-\gamma/2)}^{**}$ be the $\gamma/2$- and $(1-\gamma/2)$-quantiles, respectively, of

$$\mathcal{L}\left[\left(\hat{\phi}_T^* - \hat{\phi}_T\right)/\sqrt{\widehat{\text{var}}\left(\hat{\phi}_T^*\right)} \middle| \boldsymbol{y}_{-p+1}, \ldots, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_T; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_T\right]. \tag{5}$$

Using these quantiles, we get the studentized Hall interval

$$CI_{SH} = \left[\hat{\phi}_T - t_{(1-\gamma/2)}^{**}\sqrt{\widehat{\text{var}}(\hat{\phi}_T)}, \quad \hat{\phi}_T - t_{\gamma/2}^{**}\sqrt{\widehat{\text{var}}(\hat{\phi}_T)}\right],$$

which also has an asymptotically correct coverage probability if (5) and $\mathcal{L}[(\hat{\phi}_T - \phi)/\sqrt{\widehat{\text{var}}(\hat{\phi}_T)}]$ have identical proper limiting distributions.

A refinement of the previously considered intervals is based on the iterated bootstrap. It also is described by Hall (1992). Again, a further layer of bootstrap samples is drawn from each original bootstrap sample. Then, CI's are computed from each of the second-stage bootstrap samples and these CI's are used to estimate the actual coverage by checking how often the original estimate falls within these intervals. For example, for CI_H, let us denote the α-quantile of the second-stage bootstrap from the ith bootstrap sample by $t_\alpha^{*(2),i}$ and signify by $\mathcal{I}(\cdot)$ the indicator function. Then, the coverage probability is estimated as

$$\frac{1}{B^*} \sum_{i=1}^{B^*} \mathcal{I}\left(\hat{\phi}_T \in \left[\hat{\phi}_T^* - t_{(1-\gamma/2)}^{*(2),i}, \hat{\phi}_T^* - t_{\gamma/2}^{*(2),i}\right]\right). \tag{6}$$

If this quantity differs from the nominal coverage probability $1-\gamma$, a correction term, for example, $\hat{v}$, is determined such that (6) has correct coverage probability if $\hat{v}$ is subtracted from the lower bound and added to the upper bound of each interval. Then, the iterated CI becomes

$$CI_{IH} = \left[\hat{\phi}_T - t_{(1-\gamma/2)}^* - \hat{v}, \hat{\phi}_T - t_{\gamma/2}^* + \hat{v}\right].$$

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

This procedure also can be iterated more than once; that is, third-stage bootstrap samples can be drawn and used to estimate modifications and so on. In practice, using more than one bootstrap iteration usually will be too demanding computationally and in the example in the next section we also use just one iteration. Of course, iterated versions of CI_S and CI_{SH} could be computed in an analogous way. In the next section we use the four different CI's described in the foregoing for analyzing the impulse responses of two German monetary systems.

4. ANALYSIS OF GERMAN MONETARY SYSTEMS

Brüggemann and Wolters (1998) (BW) and Lütkepohl and Wolters (1998) (LW) consider small models for the German monetary sector to investigate the channels of monetary policy. LW use M3 as a measure of the money stock whereas BW consider a system for the narrower measure M1. In both studies, impulse responses are used to analyze the dynamic interactions of the variables in VECM's that can be represented in the form of (2). Neither study reports measures of sampling variability for the impulse responses and both conclude that the impact of the Bundesbank policy on inflation may have been quite limited because prices do not react strongly to changes in the money stock and to changes in the interest rate. In the following, we reconsider these results by checking the significance of the effects observed in the aforementioned articles. Moreover, we demonstrate the effects of using different methods for computing bootstrap CI's. We begin with a system presented by BW and then turn to one presented by LW.

4.1. M1 System

BW construct quarterly models for the period 1962:1–1989:4 and the extended period 1962:1–1996:2 using seasonal unadjusted data. In the following, we concentrate on the model version for the extended period that includes German unification in 1990 and allows for international price movements influencing domestic prices. The following variables are included in the system: $m1_t$ is the logarithm of (nominal) $M1$; y_t is the logarithm of real GNP; p_t is the logarithm of the GNP deflator, hence, $(m1 - p)_t$ is the logarithm of real $M1$ and $\Delta p_t = p_t - p_{t-1}$ is the quarterly inflation rate; R_t is a long-term interest rate ("Umlaufsrendite"); pm_t is an import price index that is treated as an unmodeled variable reflecting the openness of the German economy and capturing the effects of exchange rates. The precise data sources are provided in the Appendix. In addition, there are a number of deterministic variables in the model, such as seasonal dummies and a shift dummy, $S90q3_t$, that takes into account the level shifts in $m1_t$ and y_t due to the German unification. It is zero until 1990:2 and afterward, it has the value 1.

BW found that there is one cointegration relation between the $I(1)$ variables $m1_t$, p_t, y_t, and R_t. For the period from 1961:4 to 1996:2, they found the following long-run money demand relation [see BW, Eq. (3.4)]

$$(m1 - p)_t = 1.105y_t - 5.133R_t + 0.407S90q3_t + ec1_t. \tag{7}$$

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

TABLE 1. $M1$ system without deterministic terms, estimation period 1962:1–1996:2

Right-hand-side variables	$\Delta m1_t$	Δp_t	Δy_t	ΔR_t
$ec1_{t-1}$	−0.115 (6.6)[a]			
$\Delta m1_{t-1}$				0.036 (2.1)
$\Delta m1_{t-2}$			0.191 (3.0)	
$\Delta m1_{t-3}$			0.117 (2.0)	
Δp_{t-1}		−0.133 (2.6)		
Δp_{t-2}		0.204 (3.6)	−0.284 (3.0)	
Δp_{t-4}	0.411 (5.4)	0.538 (10.0)		
Δy_{t-2}	−0.231 (4.9)	0.082 (3.7)	−0.431 (6.4)	
Δy_{t-3}		0.113 (5.3)		
Δy_{t-4}			0.457 (7.6)	
ΔR_{t-1}	−0.884 (3.7)	0.240 (2.4)		0.192 (2.3)
Δpm_{t-4}		0.058 (2.4)		

[a] Absolute values of t-ratios in parentheses.

Here, $ec1_t$ stands for the deviations from the long-run relation. The estimated VECM of BW is given in Table 1 except for deterministic terms. The model is estimated by Zellner's seemingly unrelated regressions method. Note that the model may be viewed as a reduced form because Γ_0 is an identity matrix. Moreover, the instantaneous residual correlation is quite small and therefore no orthogonalization is needed for computing meaningful impulse responses. This model is the result of a specification procedure described in detail by BW. Their modeling procedure involves regressing the first differences of each variable on lagged differences of all the variables and the one-period lagged value of the error correction term. The specification also includes seasonal dummies and other deterministic terms. Lagged differences with insignificant coefficients are removed step by step starting with specifications that contain differenced values up to the fourth lag. Regardless of the t-values of the coefficients of the error correction term, this term was included in the regressions until all insignificant lagged differences were eliminated.

The error correction term was only eliminated if it turned out to be insignificant in the model in which the other insignificant terms were eliminated already.

Because the model is in reduced form, a fully unrestricted version with full rank error correction term may be estimated by considering the VAR form in (1) with order $p = 5$. We have used that model to compute impulse responses together with all four versions of 95% bootstrap CI's (CI_S, CI_H, CI_{SH}, CI_{IH}). The results based on 2,000 bootstrap replications are plotted in Figure 1.[1] For CI_{SH}, we used 200 bootstrap drawings for estimating $\widehat{\text{var}}(\hat{\phi}^*_T)$. Clearly, in this case the differences between the methods are not substantial. Because in most cases the CI's are almost symmetric around the estimated impulse response coefficients, it is not surprising that CI_S and CI_H are similar. Exceptions are, for instance, the responses of p, $m1$, and R to an own impulse and the response of R to an impulse in $m1$. Also, the CI_{SH} intervals are, in most cases, quite similar to CI_H. An analogous result also was obtained for other cases considered in the following. Therefore, we focus on CI_H because it has certain theoretical advantages over CI_S (see BLN) and it is much less computer intensive than CI_{SH} and CI_{IH}.

A major problem with the intervals in Figure 1 is that they are rather wide and, hence, the actual responses in the underlying system are quite uncertain if the CI's properly reflect the estimation variability. For example, based on the CI's in Figure 1, an impulse in $m1$ does not have a significant effect on the price level. Moreover, an increase in the price level does not have a significant impact on income. Thus, an impulse response analysis based on the full unrestricted reduced-form model does not give a clear indication of the relations between the variables. The results in Figure 1 also show the importance of computing CI's for the impulse responses because an interpretation that ignores the substantial estimation uncertainty may be misleading.

An improvement in the estimation precision can be expected from taking into account the restrictions imposed by BW. In Figure 2, the impulse responses and corresponding CI_H intervals that are shown are obtained for the restricted VECM presented in Table 1. Here, the cointegration vector is reestimated in each bootstrap replication. The CI_H intervals from the unrestricted VAR model are given for comparison. Obviously, taking into account the restrictions results in a substantial improvement in the precision, as expected. Now, the response of $m1$ to an impulse in the price level p has become significant and the same holds for the response of p to an impulse in $m1$, for instance. Thus, the present analysis sheds doubt on the previous interpretation from BW that the impact of changes in $m1$ on the price level may not be very strong.

Interestingly, in Figure 2, it can be seen that the impulse responses from the model with restrictions are, in most cases, within the CI's from the unrestricted model, in particular, for low lags. On the other hand, the CI's from the restricted model do not always contain the estimates of the impulse responses from the unrestricted model. Hence, estimating the impulse responses from an unrestricted model not only increases the uncertainty in the estimates but also may lead to quite different point estimates. There is more overlap between the CI's if intervals are

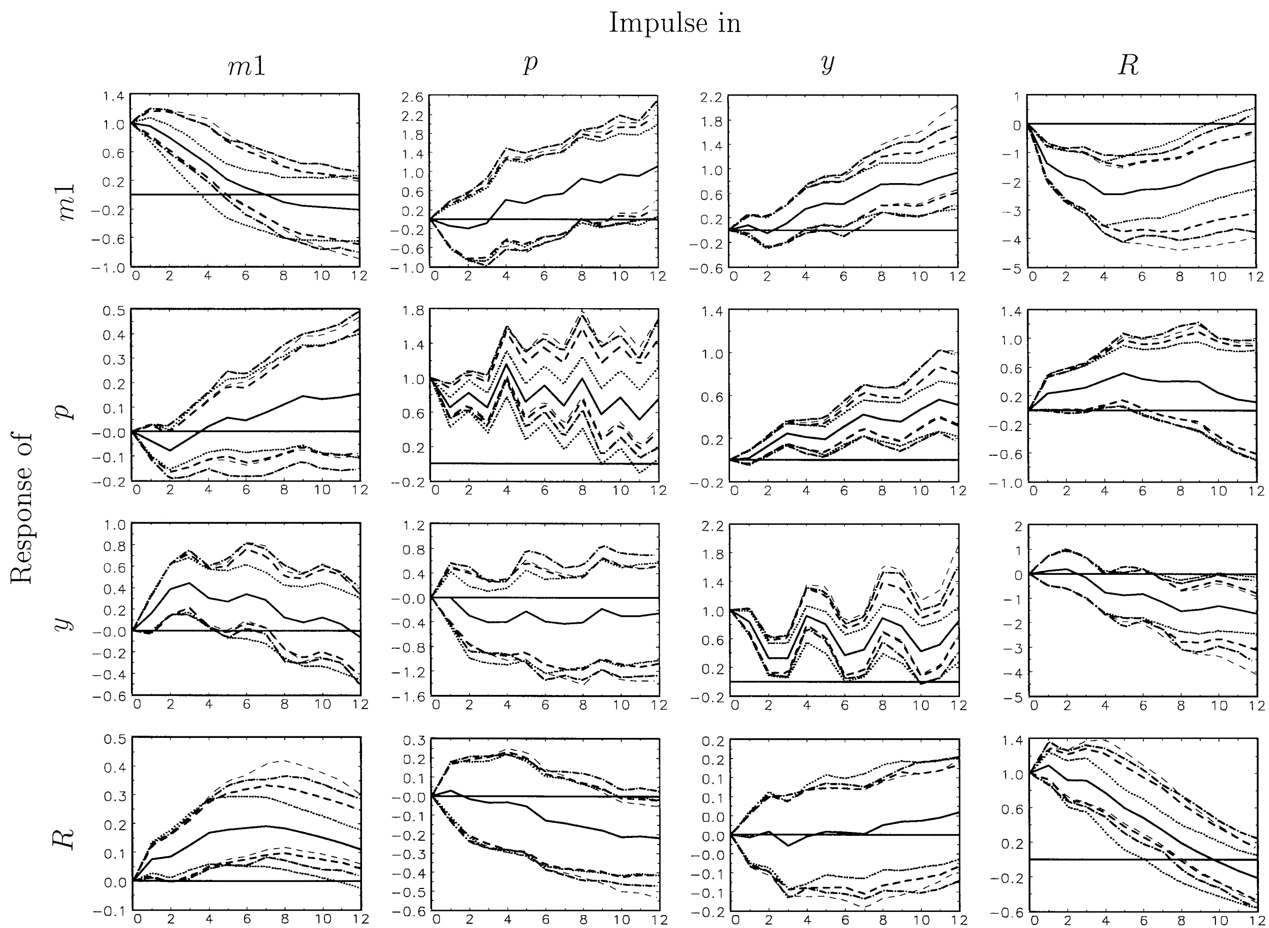

FIGURE 1. Estimated impulse responses for fully unrestricted $M1$ VAR system (solid line) with 95% CI's: CI_S (dotted lines), CI_H (bold dashed lines), CI_{SH} (light dashed lines), CI_{IH} (dots and dashes).

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

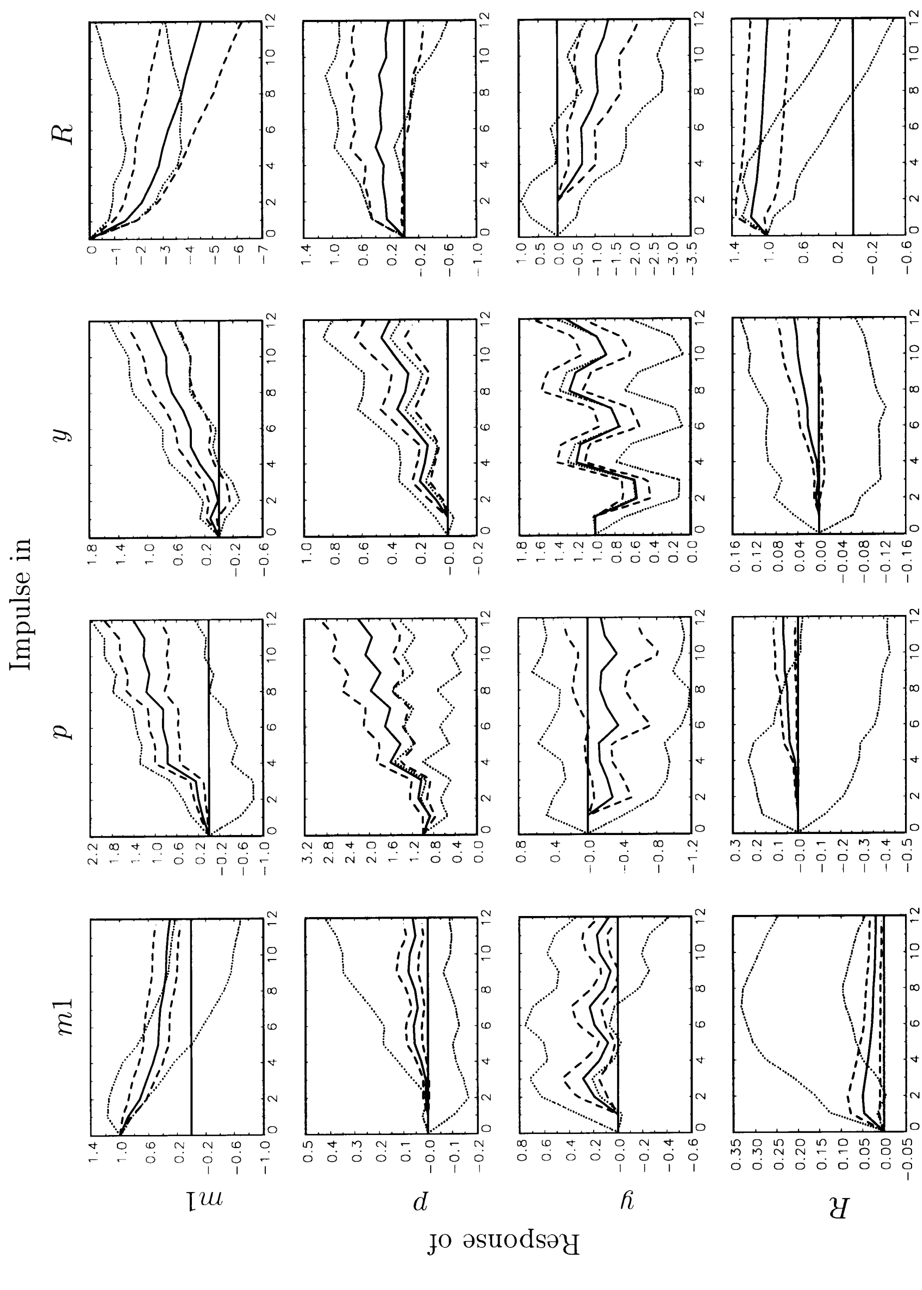

FIGURE 2. Estimated impulse responses of restricted VECM for *M1* (solid line) with 95% CI_H (dashed lines) and 95% CI_H from fully unrestricted VAR system (dotted lines).

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

computed from the restricted VECM and a VECM where only the cointegration restriction is imposed. These CI's are shown in Figure 3, where it is seen that the CI's from the less restricted model are substantially wider than the CI's from the restricted model. The long-run development of the impulse responses from both models is similar due to enforcement of the cointegration restriction. It also may be worth noting that using the bootstrap for an unrestricted model may result in singularities in the asymptotic distributions of the estimated impulse responses. This in turn may lead to strongly distorted and, hence, unreliable bootstrap CI's, as pointed out by BLN. Thus, using a restricted model is also useful for removing one source of problems for the bootstrap CI's.

The question whether to fix the estimated cointegration relation in the bootstrap or to reestimate it in each replication is addressed in Figure 4. In most cases, there is nearly no difference in the CI's. If there are differences, the CI's based on reestimated cointegration vectors tend to be larger. Of course, without a detailed analysis it is difficult to interpret this result because the reduced length intervals obtained by fixing the cointegration parameters may be the outcome of ignoring the estimation variability in the cointegration vector. Hence, it may cover up the actual estimation uncertainty that remains in the estimates. Without further knowledge of the properties of the estimates, it may be preferable to reestimate the cointegration parameters in each bootstrap replication.

4.2. M3 System

Using seasonal unadjusted data for the period 1976:1–1996:4, LW construct a quarterly model for M3. They include variables similar to those of BW in their model. In addition to the variables defined in the context of the M1 model, they use the following variables: $m3_t$ is the logarithm of (nominal) M3 and, hence, $(m3-p)_t$ is the logarithm of real M3; $(R-r)_t$ is the difference between the long-term interest rate and the own rate of M3, denoted by r_t, so that this variable represents the opportunity costs of holding M3 rather than longer-term bonds; $d_t(R-r)_t$ is identical to $(R-r)_t$ for the period 1994:3–1995:4 and is zero otherwise, and $d_t(R-r)_t$ is used to model a nonlinearity in the impact of the interest-rate differential on the demand for money in the period mentioned. The variable is treated as a member of the group of unmodeled variables in (2). Again, there are some additional deterministic variables such as seasonal dummies and dummies to take care of the unification.

LW find that the variables $(m3-p)_t$, y_t, and Δp_t are $I(1)$ and that there is one cointegration relation between these variables of the form [see LW, Eq. (3.2)]

$$(m3-p)_t = y_t - 13.50\Delta p_t + 0.14S90q3_t + ec3_t, \qquad \textbf{(8)}$$

which may be interpreted as an essential part of a long-run money demand relation. Here, $ec3_t$ represents the deviations from the long-run relation. The estimated VECM of LW is given in Table 2, where deterministic terms are excluded as in

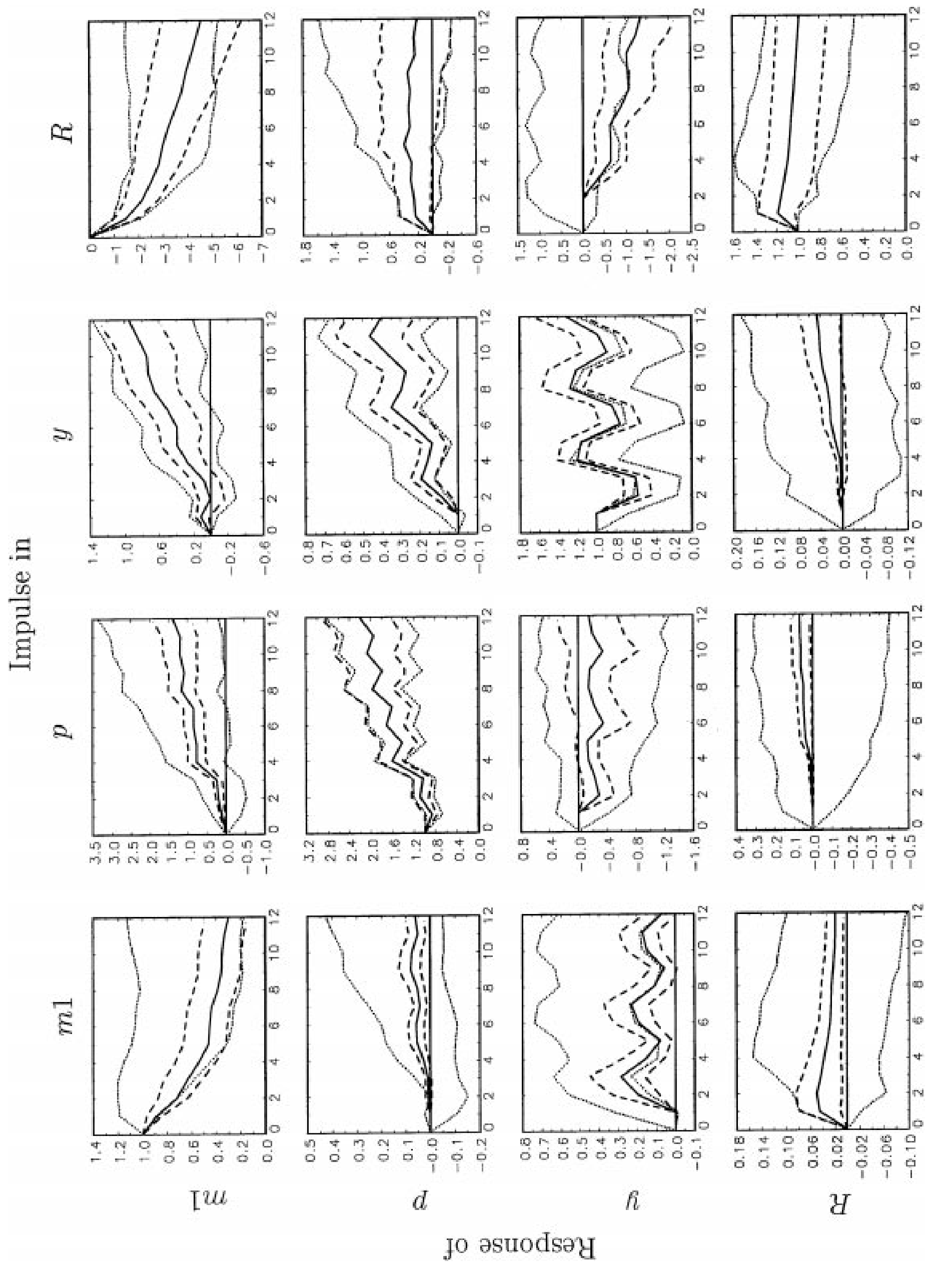

FIGURE 3. Estimated impulse responses of restricted VECM for $M1$ (solid line) with 95% CI_H (dashed lines) and 95% CI_H from unrestricted VECM (dotted lines).

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

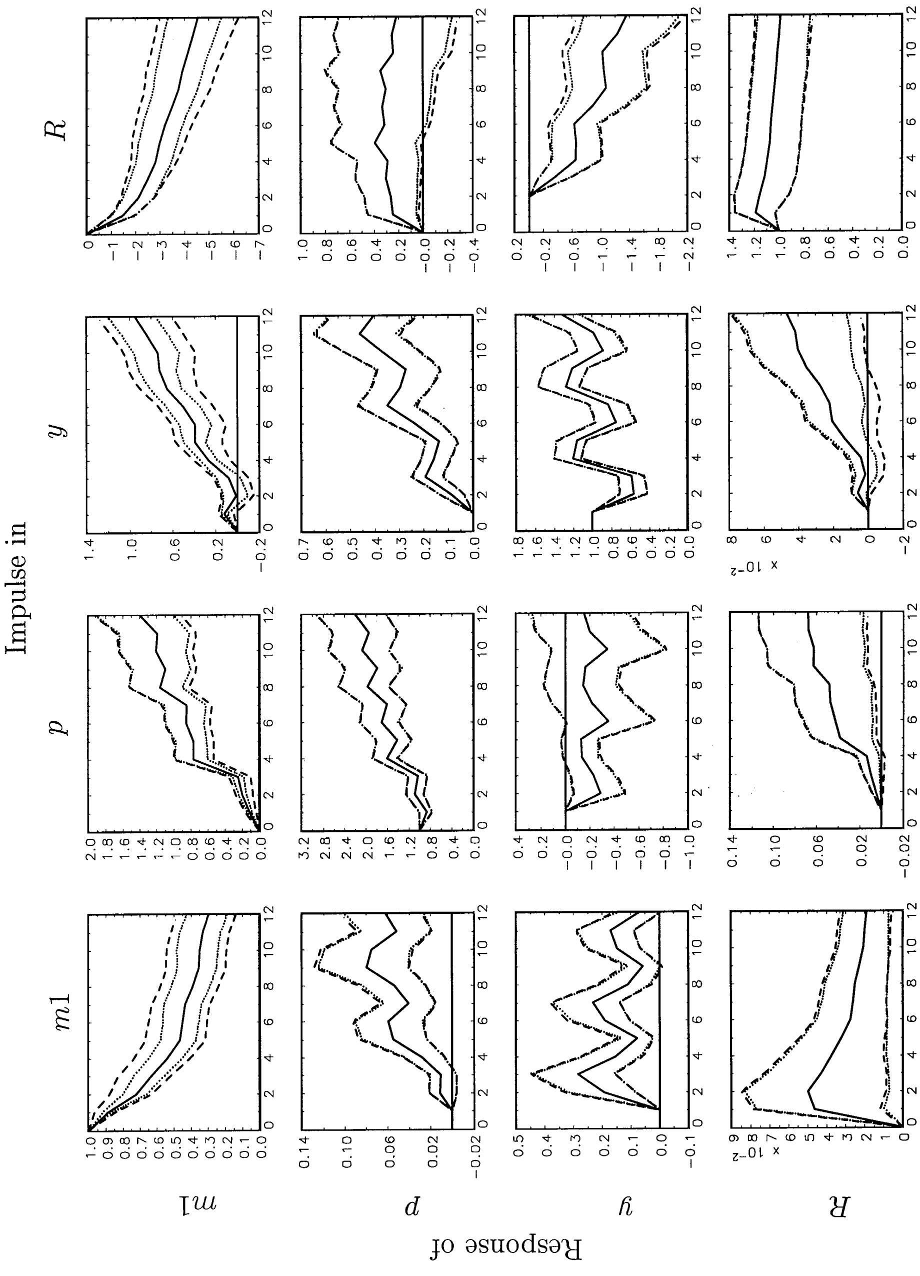

FIGURE 4. Estimated impulse responses of restricted VECM for *M*1 (solid line) with 95% CI_H when the cointegration vector is reestimated (dashed lines) or fixed (dotted lines) in every bootstrap drawing.

TABLE 2. M3 system without deterministic terms, estimation period 1976:1–1996:4

Right-hand-side variables	$\Delta(m3-p)_t$	$\Delta^2 p_t$	Δy_t	$(R-r)_t$
$ec3_{t-1}$	−0.111 (7.1)[a]		0.044 (2.3)	
$\Delta(m3-p)_{t-1}$		−0.058 (2.2)	0.269 (3.2)	
$\Delta(m3-p)_{t-2}$			0.172 (2.5)	
$\Delta(m3-p)_{t-3}$		0.089 (2.8)		
$\Delta(m3-p)_{t-4}$	−0.069 (2.0)			
$\Delta^2 p_t$	−1.262 (6.6)			
$\Delta^2 p_{t-1}$		−1.086 (11.3)		
$\Delta^2 p_{t-2}$		−1.044 (9.5)		
$\Delta^2 p_{t-3}$		−0.747 (6.9)		
$\Delta^2 p_{t-4}$	−0.251 (2.8)	−0.258 (2.7)		
Δy_{t-1}	−0.220 (5.7)		−0.323 (3.7)	
Δy_{t-3}		0.075 (2.2)		
Δy_{t-4}		0.096 (3.6)	0.243 (3.4)	
$(R-r)_{t-1}$	−0.568 (4.9)			0.836 (12.2)
$(R-r)_{t-3}$	0.427 (2.8)			
$(R-r)_{t-4}$	−0.406 (2.8)			−0.200 (2.9)
Δpm_{t-4}		0.065 (2.6)		0.055 (2.2)
$d_t(R-r)_t$	−0.430 (7.6)			

[a] Absolute values of t-ratios in parentheses.

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press

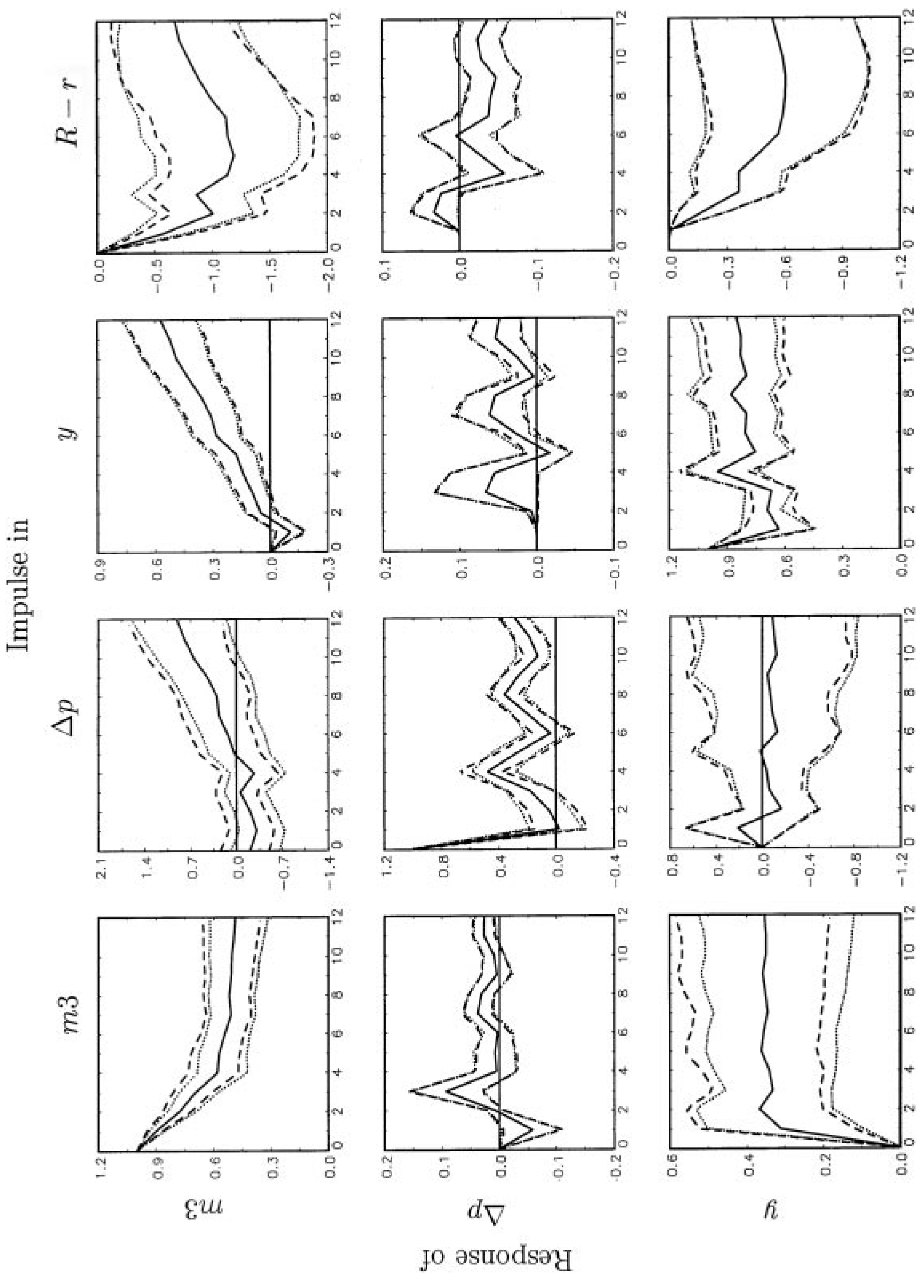

Figure 5. Estimated impulse responses of restricted VECM for $M3$ (solid line) with 95% CI's: CI_S (dotted lines), CI_H (dashed lines).

Table 1. The estimation method used is iterated three-stage least squares. The system is specified by first estimating the error correction term within a dynamic single-equation model. Then this term is treated as an additional stationary variable in specifying the system of equations. The systems specification procedure starts from a VECM, where all equations except the money equation are in reduced form. The latter equation initially contains $ec3_{t-1}$, all differences of the variables up to lag-order 4 including unlagged $\Delta^2 p_t$, $(R-r)_t$, Δy_t, and Δpm_t as well as $d_t(R-r)_t$, seasonal dummy variables, and further dummies to take care of the German unification in 1990 and other special events. The other equations also contain the lagged error correction term, four lags of the differenced endogenous variables, the unlagged Δpm_t and four lags in addition. Moreover, $d_t(R-r)_t$ as well as deterministic terms are included in the initial equations. Analogous to the M1 system, variables with insignificant coefficients then are eliminated successively according to the lowest t-values but always keeping the error correction term in each equation until the end. Then, the error correction term is excluded if its coefficient is not significant at the 5% level. More details of the specification procedure are provided by LW. Notice that the instantaneous $\Delta^2 p_t$ appears in the $\Delta(m3-p)_t$ equation and, hence, the model is a structural form in the sense that Γ_0 is not the identity matrix if the model is written in form (2). Also note that the instantaneous residual correlation is quite small, so that interpreting the residuals as impulses to specific variables is justified.

If the model is rewritten so that it looks like (1) or (2), an impulse response analysis can be carried out as described in Section 2. Since we now consider a model in structural form, we compare again CI_S and CI_H to check whether a similar result is obtained as in the reduced-form case. The impulse responses together with approximate 95% CI's are depicted in Figure 5 where the cointegration parameters are reestimated in each bootstrap replication. The impulse responses are identical to those in Figure 1 of LW. They still look a bit different because they have been scaled in a different way. The scaling in our Figure 5 is adjusted to the width of the CI's. Thus, it is less arbitrary than the scaling used by LW. It is seen in Figure 5 that the two types of CI's are again very similar. The small differences indicate that some of the underlying distributions may not be symmetric. Moreover, Figure 5 reveals that impulses in money and the interest-rate differential may have significant effects on the inflation rate. In other words, the Bundesbank's policy may have been more effective than suggested by Figure 1 of LW, thereby making apparent the importance of providing measures for the estimation uncertainty of the impulse responses as in Figure 5.

5. CONCLUSIONS

In this study we have illustrated some problems related to standard impulse response analysis in VAR models and we have suggested alternative bootstrap procedures for CI's. It has been demonstrated on the basis of two small monetary systems for Germany that it is very important to take into account that the commonly considered impulse responses are estimates and, hence, subject to some uncertainty. This estimation uncertainty has to be taken into account in the interpretation of the impulse responses. Plotting CI's together with point estimates of the impulse responses can provide a good picture of the uncertainty involved. In practice, in this context CI's often are based on bootstrap methods. We have argued that standard bootstrap CI's may be distorted and therefore may be misleading. Some alternatives from the bootstrap literature are proposed and applied for analyzing the two German monetary system examples. It is shown that the common practice of performing an impulse response analysis on the basis of a largely unrestricted model may not be very informative with respect to the actual relation of the variables because the estimation uncertainty can be substantial. Imposing restrictions on the parameters of the model on the basis of statistical criteria or a priori knowledge can lead to substantial improvements in this respect.

Note, however, that there are a number of open questions regarding the properties of the procedures used in this study. First, the asymptotic and small sample properties of bootstrap CI's in the present context are not fully clear, especially if the model contains cointegrated variables. Although there is a range of Monte Carlo studies exploring the small sample properties of estimated impulse responses, most of these studies focus on stationary VAR processes. Moreover, the underlying data generation processes are necessarily quite limited compared to the wide range of models that have been used in applied work. Hence, it is not clear whether the simulation results are generalizable to a particular model under consideration in empirical work. Second, as is common in the empirical literature, we have constructed CI's for the individual impulse response coefficients. It may be more plausible from a conceptual point of view to consider joint confidence regions for the impulse response functions because not only individual impulse response coefficients but the overall shape of some response is often of interest.

In conclusion, it is clear that there are a number of open problems surrounding impulse response analysis in the context of VAR models. Despite these problems, it is important to use the available tools for getting an impression of the uncertainty underlying any specific analysis. Therefore it is surprising that some popular software packages for dynamic econometric analysis either do not provide CI's for impulse responses (e.g., PcFiml) or provide such CI's only for simple unrestricted VAR's (e.g., Eviews), thereby complicating the interpretation of the results.

NOTE

1. The computations were performed with a GAUSS program. We have checked the sensitivity with respect to the number of bootstrap replications and found that very similar results are obtained if at least 1,000 bootstrap replications are used. Therefore, the computationally very demanding iterated bootstrap was computed with only 1,000 replications.

REFERENCES

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

Brüggemann, I. & J. Wolters (1998) Money and prices in Germany: Empirical results for 1962 to 1996. In R. Galata & H. Küchenhoff (eds.), *Econometrics in Theory and Practice*, pp. 205–225. Heidelberg: Physica.

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

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

Hendry, D.F. & G.E. Mizon (1998) Exogeneity, causality, and co-breaking in economic policy analysis of a small econometric model of money in the UK. *Empirical Economics* 23, 267–294.

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

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 Time Series Analysis*. Berlin: Springer-Verlag.

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

Lütkepohl, H. & J. Wolters (1998) A money demand system for German M3. *Empirical Economics* 23, 371–386.

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

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*, pp. 283–327. Amsterdam: North-Holland.

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

Watson, M.W. (1994) Vector autoregressions and cointegration. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, Vol. IV, pp. 2843–2915. New York: Elsevier.

APPENDIX: DATA SOURCES

Seasonal unadjusted quarterly data were used for the following variables taken from the given sources. All data refer to West Germany until 1990:2 and to the unified Germany afterward.

M1: nominal monthly values from *Monatsberichte der Deutschen Bundesbank*. The quarterly values are the values of the last month of each quarter. The variable $m1$ is ln M1.

M3: nominal monthly values from *Monatsberichte der Deutschen Bundesbank*. The quarterly values are the values of the last month of each quarter. The variable $m3$ is ln M3.

GNP: quarterly real gross national product from *Deutsches Institut für Wirtschaftsforschung, Volkswirtschaftliche Gesamtrechnung*. The variable y is ln GNP.

Price index: GNP deflator ($1991 = 100$) from *Deutsches Institut für Wirtschaftsforschung, Volkswirtschaftliche Gesamtrechnung*. The variable p is the logarithm of the price index.

Average bond rate (Umlaufsrendite) (R): monthly values from *Monatsberichte der Deutschen Bundesbank*. The quarterly value is the value of the last month of each quarter.

Own rate of M3 (r): the series was constructed from the interest rates of savings deposits (rs) and the interest rates of 3-month time deposits (rt) from *Monatsberichte der Deutschen Bundesbank* as a weighted average as follows:

$$r = \begin{cases} 0.24rt + 0.42rs & \text{for} \quad 1976{:}1\text{–}1990{:}2 \\ 0.30rt + 0.33rs & \text{for} \quad 1990{:}3\text{–}1996{:}4 \end{cases}$$

The weights are chosen according to the relative shares of the corresponding components of M3. The quarterly value is the value of the last month of each quarter.

Import price index: PM (1991 = 100) from *Deutsches Institut für Wirtschaftsforschung, Volkswirtschaftliche Gesamtrechnung*. The variable pm is the logarithm of PM.

The data can be obtained from the internet at

http://ise.wiwi.hu-berlin.de/oekonometrie/engl/data.html

https://doi.org/10.1017/S1365100501018041 Published online by Cambridge University Press