

1230

Discussion
Papers

Fundamental Problems with Nonfundamental Shocks

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2012

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

Fundamental Problems with Nonfundamental Shocks

Helmut Lütkepohl¹

Freie Universität Berlin and DIW Berlin
Mohrenstr. 58, D-10117 Berlin, Germany
email: hluetkepohl@diw.de

July 20, 2012

Abstract. Economic agents using information that is not incorporated in the econometric model is seen as a possible reason for why nonfundamental shocks are important in econometric models. Allowing for nonfundamental shocks in structural vector autoregressive (SVAR) analysis by considering moving average (MA) representations with roots in the complex unit circle is a possible response to the problem. A case is made for viewing nonfundamentalness as an omitted variables problem rather than a problem of MA roots in the unit circle. The omitted variables problem will always lurk in the background of SVAR analysis as well as other econometric studies and cannot be avoided. In SVAR analysis it is even more problematic than what the literature on nonfundamental shocks suggests. Still, SVARs can be useful tools for empirical analysis.

Key Words: Structural vector autoregression, moving average representation, vector autoregressive moving average process, impulse response analysis, factor augmented VAR, Bayesian VAR

JEL classification: C32

¹I thank Lutz Kilian and the participants of the conference on 'Nonlinear Time Series Econometrics' near Aarhus in June 2012 for comments on an earlier version of this paper.

1 Introduction

Structural vector autoregressive (SVAR) models have been used extensively for economic analysis since they were advocated by Sims (1980) as alternatives to classical econometric simultaneous equations models. Despite their popularity, a number of authors have questioned their reliability and usefulness on different grounds. For example, Cooley and LeRoy (1985) call VAR analysis atheoretical if no structural assumptions from economic theory are used in structural interpretations. Cooley and Dwyer (1998) question the robustness of the evidence from SVARs with respect to the statistical model specifications. Lütkepohl (2005, Sec. 2.3.2) and Braun and Mitnik (1993) discuss the effects of model misspecification, e.g., due to omitted variables, for impulse responses.

Another strand of critiques is based on the argument that SVARs deliver responses of the variables to fundamental shocks, while economic models may include nonfundamental shocks. Here shocks are called nonfundamental if they cannot be retrieved as forecast errors from the observed variables.² This study focusses on these types of nonfundamental shocks which are the subject of a growing number of papers and articles on structural macroeconometric modeling. Early theoretical work on nonfundamentalness goes back to Hansen and Sargent (1980, 1991). An early application illustrating the importance of nonfundamental representations was given by Lippi and Reichlin (1993). Nonfundamental representations of stochastic processes were discussed by Lippi and Reichlin (1994). More recent work on nonfundamental representations and shocks includes Giannone and Reichlin (2006), Fernández-Villaverde and Rubio-Ramírez (2006), Fernández-Villaverde, Rubio-Ramírez, Sargent and Watson (2007) and Forni, Giannone, Lippi and Reichlin (2009). Alessi, Barigozzi and Capasso (2008) review the related literature, including the more recent contributions.

The main argument in favor of nonfundamental shocks being important in economic analysis is that the econometrician does not have all the information that individual agents may have. Thus, expectations of individual agents may not be based just on past information from the variables in the empirical model. Hence, the expectation or forecast errors cannot be the residuals of the model set up by the econometrician and, thus, the shocks of interest may not be forecast errors and may be nonfundamental. I will give a critical review of some arguments used in the related literature and I will argue that

²In some of the literature, a shock is called fundamental if it has a permanent effect on variables such as output and nonfundamental if its effect is only transitory (e.g., Binswanger (2004), Lanne and Lütkepohl (2010)). This terminology is not to be confused with the terminology used in the present study.

some reactions to possible nonfundamentalness may be more plausible than others.

The two common responses to the nonfundamentalness problem are to allow for noninvertible moving average (MA) components or to view nonfundamentalness as an omitted variables problem. I will argue that the omitted variables problem is more fundamental than the nonfundamentalness problem. Solving it will also solve the nonfundamentalness problem while solving the nonfundamentalness problem without adding further variables will not solve the omitted variables problem. I will also argue that using factor models to account for larger panels of variables may not fully solve the omitted variables problem.

In the next section I formally define fundamental and nonfundamental representations of stochastic processes and the corresponding shocks. In Section 3 some arguments used in the literature on fundamental and nonfundamental representations and shocks are discussed and evaluated critically. Section 4 concludes.

The following abbreviations are used throughout: VAR for vector autoregressive or vector autoregression, SVAR for structural VAR, MA for moving average, VARMA for vector autoregressive moving average and DGP for data generation process.

2 Fundamental and Nonfundamental Moving Average Representations

It is assumed that all random variables (vectors) are from the Hilbert space $L^2(\Omega, \mathcal{F}, P)$ based on a probability space (Ω, \mathcal{F}, P) , as in Alessi et al. (2008). The symbol \mathcal{H}_t^y denotes the subspace of $L^2(\Omega, \mathcal{F}, P)$ spanned by $\{y_t, y_{t-1}, \dots\}$.

Suppose y_t is a K -dimensional stationary stochastic process with MA representation

$$y_t = \Phi(L)u_t, \quad t \in \mathbb{Z}, \quad (1)$$

where $u_t \sim (0, \Sigma_u)$ is a zero mean M -dimensional white noise process with time-invariant covariance matrix Σ_u and uncorrelated components, that is, u_t and u_s are uncorrelated for $t \neq s$. The operator

$$\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i \quad (2)$$

is a possibly infinite order $(K \times M)$ matrix power series in the lag operator L , with absolutely summable coefficient matrices Φ_i . Of course, this

setup covers cases where y_t admits VAR or VARMA representations. Note that for simplicity, the process y_t is assumed to be purely nondeterministic, that is, there are no deterministic terms in the MA representation (1). Such terms are not needed for the subsequent arguments. Moreover, I exclude processes with MA roots on the unit circle and cointegrated processes as well as other types of nonstationarities. Some of these features are important in the present context but would not change the main arguments. Hence, I prefer to exclude them for simplicity. The following definition of fundamentalness and nonfundamentalness is used (see Rozanov (1967) and Alessi et al. (2008)).

Definition. The process u_t is *y_t -fundamental* if $\mathcal{H}_t^y = \mathcal{H}_t^u$ for all $t \in \mathbb{Z}$. The process u_t is *y_t -nonfundamental* if $\mathcal{H}_t^y \subset \mathcal{H}_t^u$ and $\mathcal{H}_t^y \neq \mathcal{H}_t^u$ for any $t \in \mathbb{Z}$.

Using results of Rozanov (1967), Forni et al. (2009) state that u_t is y_t -fundamental if, for $M \leq K$, $\text{rk } \Phi(z) = M$ for all z in the complex unit circle. Thus, for $M = K$, if all roots of $\det \Phi(z)$ are outside the complex unit circle, (1) is a fundamental representation that can be inverted to a possibly infinite order VAR representation. The u_t in this case are one-step ahead forecast errors of optimal linear forecasts of y_t based on \mathcal{H}_t^y .

There are, however, equivalent nonfundamental MA representations, say

$$y_t = \Phi^*(L)u_t^*, \quad (3)$$

where $\Phi^*(z)$ has roots inside the complex unit circle. They can be obtained by post-multiplying $\Phi(L)$ by so-called Blaschke matrices, say $B(L)$, and using $u_t^* = B(L)^{-1}u_t$ (see Lippi and Reichlin (1994)). A matrix operator $B(z)$ is a Blaschke matrix if it has no poles in and on the complex unit circle and $B(z)^{-1} = B^*(\frac{1}{z})$, where the asterisk denotes the conjugate transpose. For example,

$$\begin{bmatrix} \frac{z-a}{1-\bar{a}z} & 0 \\ 0 & I_{K-1} \end{bmatrix}$$

with $|a| < 1$ is a Blaschke matrix. Here \bar{a} denotes the complex conjugate of a . An important property of Blaschke matrices is that $u_t^* = B(L)^{-1}u_t$ is white noise if u_t is white noise.

In the nonfundamental representation (3) the white noise process u_t^* is not a process consisting of forecast errors of y_t . Note, however, that the nonfundamental MA representations are equivalent to the fundamental one in the sense that $y_t = \Phi^*(L)u_t^* = \Phi(L)u_t$ and, hence, they represent the same stochastic process y_t .

Nonfundamental representations have a close relation to noncausal VARs, that is, VARs that do not have a linear representation of y_t in terms of its own lags but include future vectors y_{t+i} on the right-hand side of the equation for y_t . Such models have nonfundamental residuals and are, hence, closely related to the models considered here (see, e.g., Brockwell and Davis (1987, Chapters 3 and 11) and Lanne and Saikkonen (2010) for discussions of noncausal processes). However, in the following I focus on MA representations because they are more directly linked to impulse response analysis, which is a main tool for economic analysis. It is also worth noting that the resulting impulse responses based on nonfundamental MA representations may be quite different from the generalized impulse responses proposed by Koop, Pesaran and Potter (1996) (see also Teräsvirta, Tjøstheim and Granger (2010, Sec. 15.1)). The latter impulse responses are based on conditional expectations conditioning on past observations.

3 Fundamental versus Nonfundamental Representations

The main argument in favor of nonfundamental representations is that agents have additional information that is not known to the econometrician. Hence, agents' forecasts are not the conditional expectations obtained from the econometrician's model for y_t . Since the residuals in the fundamental representation are the forecast errors and, hence, shocks obtained from them by instantaneous linear transformations are effectively based on the econometrician's forecast errors. Thus, they may not reflect what is actually observed in practice because of the additional information agents may possess. Hence, the actual shocks of interest may have to be nonfundamental.

This argument is also used against the suitability of conventional SVAR models for analyzing economic interdependencies, because the residuals of stationary VARs with only lagged variables on the right-hand side, are forecast errors. In this setup inverting the VAR operator always leads to a fundamental MA representation.

Two alternative approaches have been proposed to cope with the nonfundamentalness problem. First, one may allow for MA representations for which the determinant has roots inside the complex unit circle and thereby open up for nonfundamental shocks. A second solution is to increase the information set, that is, the number of variables included in the set of variables under study. The latter solution has also been used to promote factor augmented VAR (FAVAR) models, which can deal with large variable sets,

and at the same time avoid the curse of dimensionality, which is problematic for unrestricted VARs given the limited length of time series available for estimation and model specification. Another approach to deal with the curse of dimensionality in VAR analysis is offered by Bayesian methods (e.g., Bańbura, Giannone and Reichlin (2010)). In the following I will first discuss the two main possibilities to respond to nonfundamentality, namely allowing for MA roots inside the unit circle and adding further information. Then I will consider FAVAR models as specific tools for dealing with more variables and also comment on Bayesian VARs.

3.1 Nonfundamentality due to MA roots in the unit circle

A fundamental MA representation can be turned into a nonfundamental one by applying the Blaschke matrices mentioned in the previous section and, conversely, nonfundamental representations can be transformed into equivalent fundamental ones. In fact, if the process is such that $\det \Phi^*(z)$ has no roots on the unit circle, there always exists an equivalent fundamental MA representation. Thus, it is not a question of the true process being fundamental or nonfundamental (e.g., Alessi et al. (2008, p. 11)). The process is always fundamental in the sense that it admits a fundamental representation, although it may be that the shocks of interest for an impulse response analysis are nonfundamental. In any case, a VAR with fundamental residuals can be fitted and used as a basis for determining any nonfundamental representation of interest by multiplying by Blaschke matrices, as described in Lippi and Reichlin (1994).

As an example, suppose that a univariate variable y_t admits an MA(1) representation,

$$y_t = u_t + \phi u_{t-1}, \tag{4}$$

where y_t -fundamentality requires that $|\phi| < 1$. An equivalent nonfundamental representation is $y_t = u_t^* + \frac{1}{\phi} u_{t-1}^*$. For the MA(1) in (4) this is actually the only basic nonfundamental representation in the sense of Lippi and Reichlin (1994) and these authors advocate using only basic ones. Clearly, $|\frac{1}{\phi}| > 1$ and, hence, the response of y_t in the next period to a nonfundamental shock can be larger than one, whereas the corresponding response to a fundamental shock is smaller than one. Of course, it is difficult to justify that a shock cannot have a response greater than one. However, that argument is not necessarily a strong argument in favour of nonfundamental shocks because a shock inducing a reaction of y_t greater than one in the next period can also

arises from adding further variables to the system under consideration, as discussed in the next subsection.

Whether it is useful or not to solve the fundamentalness problem by just allowing for MA roots in the unit circle depends on whether economic theory requires nonfundamental representations. Hansen and Sargent (1991) discuss how models with nonfundamental shocks can come about and provide examples. Obviously, allowing for nonfundamental representations in addition to fundamental ones potentially makes identification of the shocks of interest more difficult because it opens up more possibilities, except when economic theory provides fully identifying restrictions. If only fundamental shocks are considered, the shocks of interest can be obtained by linearly transforming the residuals of the reduced form VAR. If, however, also nonfundamental shocks are of potential interest, all nonfundamental representations obtained by Blaschke matrix transformations plus all admissible instantaneous linear transformations are in the competition, which may be a problem if economic theory does not deliver a full set of identifying restrictions. It is not a problem of the VAR or VARMA processes being the stochastic model for the DGP of the variables under study but just a problem of choosing the right representation for economic analysis from the portfolio of possibilities.

3.2 Nonfundamentalness due to omitted variables

Omitted variables can be another reason for nonfundamental residuals. If the econometrician considers a set of variables y_t , which does not contain all variables of importance for the economic agents, the VAR may still be a perfectly valid representation of the DGP of y_t , but due to omitted variables, may not properly reflect responses to impulses of interest. In other words, the system of interest for an economist may be $z_t = (y_t', x_t)'$, which contains y_t as a subvector plus additional variables collected in x_t . Suppose

$$z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \Theta_{yy}(L) & \Theta_{yx}(L) \\ \Theta_{xy}(L) & \Theta_{xx}(L) \end{bmatrix} \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix} = \Theta(L)u_{zt} \quad (5)$$

is a fundamental MA representation of z_t . Here the partitioning of the matrix polynomial in the lag operator $\Theta(L)$ and the residual term u_{zt} corresponds to the partitioning of z_t . Clearly, $u_{zt} = (u'_{yt}, u'_{xt})'$ generally cannot be recovered from past and present y_t , that is, \mathcal{H}_t^y is a strict subset of $\mathcal{H}_t^{u_z}$ in general and, hence, u_{zt} is not y_t -fundamental.

If a u_{yt} shock hits the system (5), the marginal responses of y_t are given by the appropriate elements of $\Theta_{yy}(L)$, which can be quite different from the impulse responses obtained from $\Phi(L)$, the MA operator from (1), for the simple reason that different shocks are considered. Of course, this insight

is not new and was pointed out long ago in the literature (see, e.g., Lütkepohl (1991, Section 2.3.2) or Giannone and Reichlin (2006) for a more recent reference). Notice, however, that both MA representations (1) and (5) are fundamental, given their respective information sets: u_t is y_t -fundamental and u_{zt} (hence, u_{yt}) is z_t -fundamental. Thus, the involved MA operators have all roots outside the complex unit circle, which is a quite different solution to the nonfundamentalness problem than the one discussed earlier.

Giannone and Reichlin (2006) make the point that the nonfundamentalness problem arises if any of the additional variables x_t is Granger-causal for y_t . As is well-known, if the MA representation is normalized such that $\Theta(0) = I$, x_t is not Granger-causal for y_t iff $\Theta_{yx}(L) \equiv 0$ (see, e.g., Lütkepohl (2005, Section 2.3.1, Proposition 2.2)). Thus, it is obvious that, if $\Theta_{yx}(L) \equiv 0$ and, hence, the MA operator in (4) is lower-block triangular, u_{yt} is y_t -fundamental. In other words, if x_t is not Granger-causal for y_t , u_{yt} can be recovered from present and past y_t . Conversely, if x_t is Granger-causal for y_t , u_{yt} is not y_t -fundamental.

To see that increasing the information set directly may be a better way of thinking about solving the fundamentalness problem than just allowing for MA roots in the unit circle, it may be worth considering a simple illustrative example. Suppose y_t and x_t are both univariate variables that are jointly generated by the bivariate MA(1)

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix} + \begin{bmatrix} \theta_{yy} & \theta_{yx} \\ \theta_{xy} & \theta_{xx} \end{bmatrix} \begin{bmatrix} u_{yt-1} \\ u_{xt-1} \end{bmatrix}. \quad (6)$$

As usual, the MA coefficients represent the responses of the variables to u_{yt} and u_{xt} shocks. For instance, θ_{yy} is the response of y_t to a u_{yt} shock one period after it occurred.

The eigenvalues of the MA coefficient matrix in (6) are

$$\lambda_{1/2} = \frac{\theta_{yy} + \theta_{xx}}{2} \pm \sqrt{\left(\frac{\theta_{yy} + \theta_{xx}}{2}\right)^2 - (\theta_{yy}\theta_{xx} - \theta_{xy}\theta_{yx})}$$

and the process is fundamental if $|\lambda_1|, |\lambda_2| < 1$. Clearly, that condition is satisfied for a large range of $\theta_{yy}, \theta_{yx}, \theta_{xy}$ and θ_{xx} values. This shows the rich set of impulse responses that are possible when the additional information in x_t is taken into account. In particular, $\theta_{yy} > 1$ is possible and, hence, a u_{yt} shock can have an effect greater than one in the period following its occurrence. For example, for $\theta_{yy} = -\theta_{xx}$ it is easy to construct cases with $\theta_{yy} > 1$ and $|\lambda_1|, |\lambda_2| < 1$. Of course, the MA coefficients depend on the properties of x_t and y_t . But given the large variety of potential x_t variables, it is difficult to exclude any feasible MA coefficient matrices a priori.

If instead of the bivariate joint process z_t , only the first component y_t is considered, then the univariate marginal process is known to admit an MA(1) representation as in (4), with ϕ being a function of θ_{yy} , θ_{yx} and the covariance matrix parameters of $(u_{yt}, u_{xt})'$ (see Sbrana and Silvestrini (2009) for the precise functional form). As mentioned earlier, in this case there is just one equivalent basic nonfundamental univariate MA(1) representation for the given marginal process y_t . As the previous discussion shows, a shock having an impact greater than one on y_t one period after its occurrence can also arise from adding further variables to the system under consideration.

Unfortunately, Granger-noncausality of x_t for y_t does not solve the problem of distorted impulse responses due to omitted variables completely because there may also be instantaneous causality if y_t and x_t are contemporaneously correlated. In that case, to get orthogonal shocks requires multiplying u_{zt} by some invertible matrix which may affect the impulse responses of y_t . Denoting by Σ_{u_z} the covariance matrix of u_{zt} and by I an identity matrix of suitable dimensions, let

$$Q = \begin{bmatrix} Q_{yy} & Q_{yx} \\ Q_{xy} & Q_{xx} \end{bmatrix}$$

be some invertible matrix partitioned according to the partitioning of z_t and such that $Q^{-1}\Sigma_{u_z}Q^{-1'} = I$, then, if $\Theta(L)$ is lower block-triangular,

$$z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \Theta_{yy}(L)Q_{yy} & \Theta_{yy}(L)Q_{yx} \\ \Theta_{xy}(L)Q_{yy} + \Theta_{xx}(L)Q_{xy} & \Theta_{xy}(L)Q_{yx} + \Theta_{xx}(L)Q_{xx} \end{bmatrix} \eta_t$$

is a possible structural fundamental MA representation with orthogonal shocks $\eta_t = Q^{-1}u_{zt}$. Hence, all shocks in η_t may have nonzero effects on y_t although x_t is not Granger-causal for y_t . Thus, fundamentalness is necessary but not sufficient for informative impulse responses when important variables are omitted from the analysis. Hence, even if x_t is not Granger-causal for y_t , it may have to be considered in a proper, undistorted impulse response analysis. This problem obviously cannot be solved by allowing for noninvertible MA representations. These considerations show that nonfundamentalness is only part of the problem caused by omitted variables.

The bottom line of this discussion is that there are two ways for overcoming the problem that the fundamental representation of the marginal process of y_t may not adequately reflect the responses to shocks. One is based on remaining within the set of variables y_t and considering nonfundamental MA representations of y_t , while the second possibility is to increase the dimension of the process considered. The former possibility can be quite restrictive compared to the second possibility because it does not account for instantaneous

relations between included and omitted variables. When the basic argument for nonfundamentalness is that the econometrician has not accounted for all relevant information, it makes sense to add the missing information to the system to be considered. Which solution is favored in a particular case may also depend on the underlying economic theory. In any case, it does not mean that (causal) VARs are unsuitable models for the DGP of the set of variables of interest. It shows, however, a potential limitation of fitting small dimensional VARs.

It may also be worth reflecting on the nature of the shocks. In SVAR analyses typical shocks of interest are demand, supply, monetary or technology shocks. Suppose a technology shock is of interest. Clearly in a small VAR, where real variables such as output and consumption are not broken down in sectors, a technology shock can only be thought of as some kind of average technology shock and the impulse responses are the responses to such an average shock. In practice, average shocks rarely (if ever) hit the system but a technological innovation may occur in some sector and from there it may be transmitted to other sectors and eventually it may affect the whole economy. Tracing the transmission from one sector to the system may not only give a different picture than tracing an average shock but also requires that sectoral variables are contained in the system under consideration.

If one looks only at the marginal univariate process $y_t = u_t + \phi u_{t-1}$ of the first component of the bivariate example process (6), then tracing a u_t shock corresponds to tracing an average shock to variable y_t . A more detailed, possibly more disaggregated picture is obtained by tracing u_{yt} or u_{xt} shocks in the bivariate process (6). Actually there has been some literature considering effects of shocks to univariate variables to investigate the persistence of time series variables (e.g., Nelson and Plosser (1982), Kilian and Zha (2002)).

As another example, consider a monetary policy shock. For such shocks it may be more natural to think of them as hitting all parts of the economy at once because markets react quickly to central bank decisions. Even so, there is evidence that a change in a short-term interest rate has only a gradual impact on longer-term rates, as the literature on the pass-through of changes in interest rates shows (e.g., de Bondt (2005)). Taking such effects into account in a model which includes a range of interest rates over the whole term structure may give a different impression than a small model with only one (short-term) interest rate. Which kind of model should be considered depends also on which types of questions one is interested in. If interest is in the main responses to some average shock, looking at a small model may be quite reasonable although one has to be aware that what the resulting impulse responses show may be quite different from what one sees when a specific shock hits.

3.3 Avoiding nonfundamentality by using factor augmented or large Bayesian VARs

These and other considerations have led to another argument against using VARs and a possibility for overcoming nonfundamentality that has been suggested in the literature. It has been argued that standard VARs typically include only a small number of variables because otherwise the estimation and specification uncertainty becomes so large that the models become uninformative with respect to the relations between the variables involved or degrees of freedom for estimation are exhausted. In economics there is typically a large number of relevant variables that may indeed be too large relative to the sample size and may make it impossible to estimate an unrestricted VAR. There are a number of proposals for overcoming this problem. For example, one may impose zero restrictions on the VAR coefficient matrices as in subset models (e.g., Lütkepohl (2005, Section 5.2)) or consider so-called global VARs (e.g., Pesaran, Schuermann and Weiner (2004)) or one may impose shrinkage restrictions on the parameters (e.g., Litterman (1986), De Mol, Giannone and Reichlin (2008), Bańbura et al. (2010)).

Another approach which enjoys increasing popularity is the use of FAVAR models. Factor models for reducing the dimensionality of multiple time series models have a long tradition in the econometrics literature (see, e.g., Sargent and Sims (1977), Geweke (1977), Engle and Watson (1981), Forni, Hallin, Lippi and Reichlin (2000), Forni and Lippi (2001), Stock and Watson (2002a, b, 2005), Bai (2003), Bai and Ng (2002)). In applied work they have been used more extensively only lately since suitable databases have become available (see the overview of applications in Breitung and Eickmeier (2005)). The idea is to aggregate the information in a large number of time series in a small number of factor time series and use them in a VAR model instead of the large set of original series.

I use the following simplified setup of Forni et al. (2009), which is general enough to discuss the main points of interest. It is similar to the setups used by a number of other authors in the related literature. The observed variables of interest are again denoted by y_t . They are assumed to be generated as

$$\begin{aligned} y_t &= \Lambda f_t + \xi_t, \\ f_t &= \Psi(L)v_t, \end{aligned} \quad t \in \mathbb{Z}, \quad (7)$$

where f_t is a q -dimensional vector of common factors, ξ_t is a K -dimensional stationary stochastic process of idiosyncratic components, Λ is a $(K \times q)$ dimensional matrix of factor loadings, $\Psi(L)$ is a $(q \times q)$ dimensional matrix polynomial in the lag operator and v_t is a white noise error term. Actually Forni et al. (2009) allow the dimension K of y_t to go to infinity and they

also allow the number of shocks v_t driving the factors to be smaller than the number of common factors. I do not consider those generalizations because the main arguments can be discussed in the present framework.

The setup (7) implies that

$$y_t = C(L)v_t + \xi_t,$$

where $C(L) = \Lambda\Psi(L)$ is $(K \times q)$ and y_t is essentially driven by a small(er) number of q shocks v_t . Forni et al. (2009) argue that $w_t = C(L)v_t$ is fundamental under rather mild conditions because the requirement for fundamentalness ($\text{rk } C(z) = q$) is satisfied under realistic conditions. To me these arguments are not fully convincing, however.

First of all, $y_t \neq w_t = C(L)v_t$ in general. Hence, the assumption that $w_t = C(L)v_t$ is the relevant set of variables to look at means that the actual variables of interest are latent variables that are unobservable but could, of course, still be the relevant variables from an economic theory point of view. Given that economic models are often formulated without error terms, this assumption may be attractive to some researchers. Since $K > q$, there must be fixed linear relations between the variables w_t , which I do not find particularly appealing for observed economic variables. Of course, there are accounting identities that make some variables linearly dependent by definition. Apart from such linear dependencies, economic variables are better viewed as stochastic, meaning that only variables are included in y_t for which ξ_t has a nonsingular covariance matrix. In that case the MA representation again becomes $y_t = \Phi(L)u_t$ and the fundamentalness condition is that the roots of $\det \Phi(z)$ are outside the unit circle. If there are reasons that this cannot be assumed, the situation is as before and factors do not help, except, of course, in the sense that they ensure that a lot more information is included than in small scale conventional VARs.

It may be worth pointing out that if the latent variables w_t respond to v_t shocks as prescribed by $C(L)$, this does not mean that the y_t responses, that is, the responses of the observed counterparts that are supposedly of interest, are the same. Assuming that interest is only in the marginal responses of the y_t variables to v_t innovations may be plausible in some situations. Disentangling the effects of v_t shocks and other shocks hitting at the same time requires rather stringent orthogonality restrictions for v_t and ξ_t . Again the question is whether the analyst is interested in what is actually observed after a shock hits or what might happen in some latent background system.

To see the difference, consider a univariate MA(1), $w_t = v_t + cv_{t-1}$ with $|c| < 1$. Thus,

$$y_t = v_t + cv_{t-1} + \xi_t = u_t + \phi u_{t-1}.$$

Clearly, a fundamental v_t shock will have the effect c on w_t and, hence, on y_t after one period and this is different from ϕ , the corresponding effect of a u_t innovation. Assuming that ξ_t is white noise and denoting by σ_v^2 , σ_ξ^2 and σ_u^2 the variances of v_t , ξ_t and u_t , respectively, gives

$$E(y_t^2) = (1 + c^2)\sigma_v^2 + \sigma_\xi^2 = (1 + \phi^2)\sigma_u^2$$

and

$$E(y_t y_{t-1}) = c\sigma_v^2 = \phi\sigma_u^2.$$

Apparently there is an identification problem here, that is, there are degrees of freedom in choosing c , σ_v^2 and σ_ξ^2 for given ϕ and σ_u^2 . If economics is silent about how to split up y_t in w_t and ξ_t , some statistical normalization is required which in the end determines the effect c . That effect is then interpreted by economists. Without much economic theory there is an element of arbitrariness here.

Things may not look quite so bad in FAVAR models because the split-up is conditional on a small number of shocks in w_t . Clearly, in this case the choice of the number of factors becomes critical. In any case, the factor structure setup is a statistical identification or normalization device that has little to do with economics per se, although it may be good for forecasting, for example. The fact that observed economic variables are not exactly linearly related (excluding accounting identities) means effectively that there cannot be fewer shocks than variables. Of course, it is quite possible that there are fewer *shocks of interest* than variables. The factor structure is just a convenient simplification to avoid degrees of freedom problems in estimating large-dimensional models. It has little to do with economics. If economics tells us that there are nonfundamental shocks, they also have to be accommodated in factor models. Whether or not nonfundamental shocks are a useful device in economic models is open for discussion. Such a discussion is beyond the scope of this study.

A problem with FAVAR models more generally is that they impose a rather stringent structure on the MA operator $\Phi(L)$. If they are merely viewed as approximations to the true DGP, which are necessary to include a lot of variables, it is not at all clear that such restricted MA operators provide the true responses to shocks. Thus, factor models are a possible response to the omitted variables and curse of dimensionality problems in conventional VARs but may not fully solve them. In any case, the user of such models should be aware that the related impulse responses are only approximately valid and the approximation may be good or bad. The factors summarize the information in a large set of variables in a specific way. Whether or not this

summary is better than simple aggregation of, say sectoral, disaggregated information is not at all clear. Neither kind of information aggregation takes into account the economic relations between the disaggregated components. Hence, it is not clear that factor models provide the superior framework for impulse response analysis.

The same problem is encountered in many large-scale Bayesian VARs, where the prior is typically not derived from economic theory. For example, Bańbura et al. (2010) use the classical Litterman or Minnesota prior as basis for their analysis, which was derived by considering statistical properties of economic variables. Priors are used in large-scale Bayesian VARs to compensate for degrees-of-freedom limitations. Hence, they are bound to distort the parameter estimates and also the estimated impulse responses. Although the distortions may not be the same as in FAVAR models it is clear that Bayesian VAR models do not resolve the problem of inducing distortions to the impulse responses.

Given these problems, it does not appear to be a good idea to interpret the actual values of impulse responses too literally. They can tell us something about the direction of reactions of variables to average shocks or more generally the shape of the response with some luck. It is not likely that we can infer the precise quantitative responses to specific shocks from them. This, of course, is a common situation in statistics: it tells us about average tendencies, not specific realizations of random variables. This is not to say that economists should ignore impulse responses. They do provide information on economic relations which has to be interpreted in the right way. Doing so is easier when the limitations of impulse response analysis are better understood. Contributing to that objective is the purpose of these reflections.

4 Conclusions

One argument against using conventional SVARs for macroeconomic analysis is that they deliver fundamental shocks. These are shocks that can be recovered from past and present observed variables. In turn, shocks that are not recoverable from present and past observations are called nonfundamental. The fact that economic agents use additional information in decision making that is not available to, or at least not used by, the econometricians in setting up the model is a typical argument in favor of nonfundamental shocks. Two approaches have been proposed to address the nonfundamentality problem. First, in fundamental MA representations all roots of the determinantal polynomial obtained from the MA operator are outside the

complex unit circle. Thus, allowing for roots inside the unit circle produces nonfundamental representations and shocks. The second common approach to the nonfundamentalness problem is increasing the information set, that is, the variables considered in the VAR. Both approaches are plausible solutions to the nonfundamentalness problem but also have their own limitations.

Allowing for MA roots in the unit circle is plausible if the underlying economic model has this feature. It is not clear, however, that the full range of impulse responses that may be observed in practice can really be generated by just allowing for nonfundamental MA representations. Moreover, as I have argued, the identification of shocks of interest becomes much harder than in a purely fundamental framework, unless economic theory provides a full set of identifying restrictions. Unfortunately, the latter condition is rarely satisfied in practice.

Viewing nonfundamentalness as an omitted variables problem is attractive in view of the argument for how nonfundamentalness comes about. If it is due to information not available or used by the econometrician, adding the missing information to the model is a natural response to the problem. Unfortunately, adding all information that may be important in agents' decision making may lead to a large number of variables and to the well-known curse of dimensionality problem. The latter problem has been tackled in different ways. One proposal is to use FAVAR models. I have argued that, while these models are a plausible response to omitted variables and they are useful for integrating large panels of variables, they may not provide realistic impulse responses.

In summary, the growing literature on nonfundamental shocks in SVAR models has proposed plausible solutions that are beset with their own problems and are not likely to provide satisfactory ways out of the dilemma that economic agents have large information sets at their disposal in making decisions while VARs can realistically only handle a very limited number of variables. In the end, any proposal for overcoming the curse of dimensionality, which results from including large numbers of variables, can be criticized on the grounds that it may induce distortions in impulse responses. The same holds for large-scale Bayesian VAR models that have been proposed as an alternative way for including the information from many variables in a VAR model.

Indeed, nonfundamentalness is only one aspect that makes impulse response analysis problematic. Lütkepohl (2005, Section 2.3.2) points out other problem areas as well. For example, aggregation (contemporaneous as well as temporal) can distort impulse responses. Starting from continuous time processes the temporal aggregation problem was also discussed by Hansen and Sargent (1991). Moreover, seasonal adjustment, trend adjust-

ment or measurement errors in variables may induce distortions of impulse responses.

I do not think, however, that all these problems provide strong arguments against using VARs for macroeconometric analysis. Indeed, similar arguments could be raised against most kinds of econometric analysis. Still, for a deep understanding of the mechanisms behind economic systems, it is fundamental to look at the evidence contained in the data. For properly interpreting the evidence from a VAR analysis it is important to understand the limitations of the model. These thoughts are meant to contribute to a better understanding. In my view, even the fact that adding a new variable to a small VAR changes or does not change the impulse responses is an important and potentially interesting finding that may help us understand what is going on in a particular system. Of course, given all the possible distortions, one may want to be careful in interpreting the impulse responses. The actual numbers will, indeed, not reflect precisely what actually happens when a shock hits a system. Hence, one may not want to interpret every wiggle in an impulse response function as an important empirical finding. Still, not trying to learn anything from SVAR impulse responses would be the wrong reaction to these problems.

References

- Alessi, L., Barigozzi, M. and Capasso, M. (2008). A review of nonfundamentalness and identification in structural VAR models, *Working Paper 922*, European Central Bank, Frankfurt.
- Bai, J. (2003). Inferential theory for factor models of large dimensions, *Econometrica* **71**: 135–171.
- Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models, *Econometrica* **70**: 191–221.
- Bańbura, M., Giannone, D. and Reichlin, L. (2010). Large Bayesian vector autoregressions, *Journal of Applied Econometrics* **25**: 71–92.
- Binswanger, M. (2004). How important are fundamentals? - Evidence from a structural VAR model for the stock markets in the US, Japan and Europe, *Journal of International Financial Markets, Institutions and Money* **14**: 185–201.
- Braun, P. A. and Mittnik, S. (1993). Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions, *Journal of Econometrics* **59**: 319–341.

- Breitung, J. and Eickmeier, S. (2005). Dynamic factor models, *Discussion Paper 38*, Deutsche Bundesbank.
- Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*, Springer-Verlag, New York.
- Cooley, T. F. and Dwyer, M. (1998). Business cycle analysis without much theory: A look at structural VARs, *Journal of Econometrics* **83**: 58–88.
- Cooley, T. F. and LeRoy, S. F. (1985). Atheoretical macroeconometrics: A critique, *Journal of Monetary Economics* **16**: 283–308.
- de Bondt, G. J. (2005). Interest rate pass-through: Empirical results for the euro area, *German Economic Review* **6**: 37–78.
- De Mol, C., Giannone, D. and Reichlin, L. (2008). Forecasting using a large number of predictors - is Bayesian regression a valid alternative to principal components?, *Journal of Econometrics* **146**: 318–328.
- Engle, R. F. and Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates, *Journal of the American Statistical Association* **76**: 774–781.
- Fernández-Villaverde, J. and Rubio-Ramírez, J. F. (2006). Economic and VAR shocks: What can go wrong?, *Journal of the European Economic Association* **4**: 466–474.
- Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J. and Watson, M. W. (2007). ABCs (and Ds) of understanding VARs, *American Economic Review* **97**: 1021–1026.
- Forni, M., Giannone, D., Lippi, M. and Reichlin, L. (2009). Opening the black box: Structural factor models with large cross sections, *Econometric Theory* **25**: 1319–1347.
- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000). The generalized dynamic factor model: Identification and estimation, *Review of Economics and Statistics* **82**: 540–552.
- Forni, M. and Lippi, M. (2001). The generalized dynamic factor model: Representation theory, *Econometric Theory* **17**: 1113–1141.
- Geweke, J. (1977). The dynamic factor analysis of economic time-series models, in D. J. Aigner and A. S. Goldberger (eds), *Latent Variables in Socio-Economic Models*, North-Holland, New York, pp. 365–383.
- Giannone, D. and Reichlin, L. (2006). Does information help recovering structural shocks from past observations?, *Journal of the European Economic Association* **4**: 455–465.

- Hansen, L. P. and Sargent, T. J. (1980). Formulating and estimating dynamic linear rational expectations models, *Journal of Economic Dynamics and Control* **2**: 7–46.
- Hansen, L. P. and Sargent, T. J. (1991). Two difficulties in interpreting vector autoregressions, in L. P. Hansen and T. J. Sargent (eds), *Rational Expectations Econometrics*, Westview Press, Boulder, pp. 77–120.
- Kilian, L. and Zha, T. (2002). Quantifying the uncertainty about the half-life of deviations from PPP, *Journal of Applied Econometrics* **17**: 107–125.
- Koop, G., Pesaran, M. H. and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics* **74**: 119–147.
- Lanne, M. and Lütkepohl, H. (2010). Structural vector autoregressions with non-normal residuals, *Journal of Business & Economic Statistics* **28**: 159–168.
- Lanne, M. and Saikkonen, P. (2010). Noncausal vector autoregression, *Working paper*, University of Helsinki.
- Lippi, M. and Reichlin, L. (1993). The dynamic effects of demand and supply disturbances: Comment, *American Economic Review* **83**: 244–252.
- Lippi, M. and Reichlin, L. (1994). VAR analysis, nonfundamental representations, Blaschke matrices, *Journal of Econometrics* **63**: 307–325.
- Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions - five years of experience, *Journal of Business & Economic Statistics* **4**: 25–38.
- Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*, Springer Verlag, Berlin.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of Monetary Economics* **10**: 139–162.
- Pesaran, M. H., Schuermann, T. and Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconomic model, *Journal of Business & Economic Statistics* **22**: 129–162.
- Rozanov, Y. A. (1967). *Stationary Random Processes*, Holden-Day, San Francisco.
- Sargent, T. J. and Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory, in C. A. Sims (ed.), *New Methods in Business Cycle Research: Proceedings from a Conference*, Federal Reserve Bank of Minneapolis, Minneapolis, pp. 45–109.

- Sbrana, G. and Silvestrini, A. (2009). What do we know about comparing aggregate and disaggregate forecasts?, *Discussion Paper 2009/20*, CORE, Université Catholique de Louvain, Belgium.
- Sims, C. A. (1980). Macroeconomics and reality, *Econometrica* **48**: 1–48.
- Stock, J. H. and Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association* **97**: 1167–1179.
- Stock, J. H. and Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes, *Journal of Business & Economic Statistics* **20**: 147–162.
- Stock, J. H. and Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis, *manuscript*, Princeton University.
- Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J. (2010). *Modelling Nonlinear Economic Time Series*, Oxford University Press, Oxford.