Quantitative Macroeconomic Modeling with Structural Vector Autoregressions – An EViews Implementation

S. Ouliaris[1], A.R. Pagan[2] and J. Restrepo[3]

September 19, 2016

[1]sam.ouliaris@gmail.com
[2]apagan@unimelb.edu.au
[3]jrestrepo@imf.org

Preface

This book began as a series of lectures given by the second author at the International Monetary Fund as part of the Internal Economics Training program conducted by the Institute for Capacity Development. They continued for the years from 2011-2015 and were gradually adapted to describe the methods available for the analysis of quantitative macroeconomic systems with the Structural Vector Autoregression approach. A choice had to be made about the computer package that would be used to perform the quantitative work and EViews was eventually selected because of its popularity amongst IMF staff and central bankers more generally. Although the methodology developed in this book extends to other packages such as Stata, it was decided to illustrate the methods with EViews.

Because the book developed out of a set of lectures we would wish to thank the many IMF staff and country officials who participated in the courses and whose reactions were important in allowing us to decide on what should be emphasised and what might be treated more lightly. The courses were exceptionally well organized by Luz Minaya and Maria (Didi) Jones. Versions of the course were also given at the Bank of England and the Reserve Bank of Australia. We also need to thank Gareth Thomas and Glenn Sueyoshi at IHS Global Inc for their help in understanding some of the functions of the EViews package and for providing some new options in EViews 9.5 that we found important for implementing the methods of this book.

Finally, on a personal level Adrian would like to dedicate the book to Janet who had to endure the many hours of its construction and execution.

Contents

List of Figures

List of Tables

Chapter 1

An Overview of Macro-econometric System Modeling

Capturing macro-economic data involves formulating a system that describes the joint behavior of a number of aggregate variables. Early work in econometrics tended to focus upon either the modeling of a single variable or the small set of variables such as the price and quantity of some commodity which would describe a market. Although there was much theoretical work available about the inter-relationships between aggregates such as output, money and prices, Tinbergen (1936, 1939) seems to have been the first to think about capturing these quantitatively through the specification of a system of equations, followed by the estimation of their parameters. A modeling methodology then developed which centered upon a set of *reduced form equations* for summarizing the data and a set of *structural form equations* for interpreting the data. Variables were classified as whether they were endogenous (determined within the macro-economic system) or exogenous (roughly, determined outside of the system). The reduced form equations related the endogenous to the exogenous variables, while the structural equations aimed to capture the relationships that the endogenous variables bore to one another as well as to some of the exogenous variables. Often structural equations were thought of as describing the type of decision rules that were familiar from economic theory, e.g. consumers demanded a certain quantity of aggregate output based on the aggregate price level as well as how liquid they were, with the latter being measured by real money holdings.

The development of the concepts of reduced-form and structural equations led to the question of the relationship between them. To be more concrete about this, think of the endogenous variables z_{1t} and z_{2t} as being the money stock and the interest rate respectively, and let x_{1t} and x_{2t} be exogenous variables. Then we might write down two structural equations describing the demand and supply of money as

$$z_{1t} = a_{12}^0 z_{2t} + \gamma_{11}^0 x_{1t} + \gamma_{12}^0 x_{2t} + u_{1t} \tag{1.1}$$
$$z_{2t} = a_{21}^0 z_{1t} + \gamma_{22}^0 x_{2t} + \gamma_{21}^0 x_{1t} + u_{2t}, \tag{1.2}$$

with the errors u_{jt} treated as random variables. For a long time in econometrics these errors were assumed to have a zero expectation, to be normally distributed with a covariance matrix Ω_S, and for u_{jt} to have no correlation with past u_{kt}. The reduced form of this system would have the form

$$z_{1t} = \pi_{11} x_{1t} + \pi_{12} x_{2t} + e_{1t} \tag{1.3}$$
$$z_{2t} = \pi_{21} x_{1t} + \pi_{22} x_{2t} + e_{2t}, \tag{1.4}$$

where $e_{1t} = \delta_{11} u_{1t} + \delta_{12} u_{2t}$, and the δ coefficients are weighted averages of the structural parameters a_{ij}^0, γ_{ij}^0. The coefficients π_{ij} are also functions of the structural coefficients a_{ij}^0, γ_{ij}^0 and e_t would have zero mean and be normally distributed with covariance matrix Ω_R. Again there was no correlation between e_t and past values. Therefore, because the parameters of any structural form such as (1.1) - (1.2) were reflected in the reduced form parameters, it was natural to ask whether one could recover unique values of the former from the latter (the π_{ij}). This was the *structural identification* problem.

The reduced form contained all the information that was in the data and described it through the parameters π_{ij}. Consequently, it was clear that this created an upper limit to the number of parameters that could be estimated (identified) in the structural form. In the example above there are seven parameters in the reduced form - four π_{ij} and three in Ω_R - so the nine in (1.1) - (1.2) cannot be identified. The conclusion was that the structural equations needed to be simplified. One way to perform this simplification was to exclude enough of the endogenous or exogenous variables from each of the equations. This lead to rank and order conditions which described the number and type of variables that needed to be excluded.

Now owing to the exogenous variables present in (1.3) - (1.4) the parameters of the reduced form could be estimated via regression. Hence, it was not necessary to know what the structural relations were in order to summarize the data. This then raised a second problem: there could be many structural forms which would be compatible with a given reduced form, i.e. it might not be possible to find a *unique structural model*. As an example of this suppose we looked at two possibilities. First set $\gamma_{11}^0 = 0, \gamma_{22}^0 = 0$, i.e. exclude x_{1t} from (1.1) and x_{2t} from (1.2). Then consider the alternative $\gamma_{12}^0 = 0, \gamma_{21}^0 = 0$. In both cases there are the same number of structural parameters as in the reduced form. Consequently, there is no way to choose between these two models because π_{ij} and Ω_R can be found regardless of which is the correct structural model. These models are said to be *observationally equivalent* and the fact that there is a range of models could be termed the issue of *model identification*. The structural identification solution only took a *given* structural form and then asked about determining its parameters uniquely from the reduced form. It did not ask if there was more than one structure that was compatible with the same reduced form. Model identification was generally not dealt with in any great detail and structural identification became the most studied issue. Preston (1978) is one of the few to make this distinction.

Although the study of identification proceeded in a formal way it was essentially driven by what was termed the "regression problem". This arose since running a regression like (1.1) would give inappropriate answers since the regression assumption was that z_{2t} was uncorrelated with u_{1t}, and (1.4) showed that this was generally not true. In order to obtain good estimates of the structural parameters it was necessary to "replace" the right hand side (RHS) endogenous variables in (1.1) and (1.2) with a measured quantity that could be used in regression, and these were termed *instruments*. Further analysis showed that an instrument had to be correlated with the endogenous variable it was instrumenting and uncorrelated with the error term of the regression being analyzed. Exogenous variables were potential instruments as they were assumed uncorrelated with the errors. Therefore the question that needed to be settled was whether they were related to the variable they instrumented, i.e. whether the relevant π_{ij} were non-zero? Of course one couldn't replace z_{2t} in (1.1) with either x_{1t} or x_{2t}, as these variables were already in the regression and there would be collinearity. One needed some exclusion restrictions associated with the exogenous variables such as $\gamma_{12}^0 = 0$, and then z_{2t} might be "replaced" by x_{2t}. Reasoning such as this led to conclusions in respect to identification that coincided with the order condition.[1] The rank condition revolved around further ensuring that in reality there was a relation between a variable like z_{2t} and x_{2t}, i.e. $\pi_{22} \neq 0$.

Tinbergen and others realized that the relations between the variables needed to be made dynamic, as the responses of the endogenous variables to changes in the exogenous variables did not occur instantaneously, but slowly over time. This had two consequences. One was that lagged values such as z_{jt-1} might be expected to appear in (1.1) and (1.2), and these became known as *predetermined variables*. In many formulations of structural equations it might be expected that z_{jt-1} would appear in the $j'th$ equation but not others, so that seemed to provide a "free good", in the sense of leading to many instruments that could be excluded from the structural equations. Accordingly, it might be expected that identification would always hold. It also led to the idea of constructing dynamic multipliers which described the responses of the endogenous variables as time elapsed after some stimulus.

The recognition of lags in responses also led to the idea that one might think of variables as being determined sequentially rather that simultaneously i.e. $a_{12}^0 = 0$. Nevertheless it was recognized that a sequential structure did not fully solve the regression problem. If $a_{12}^0 = 0$ it did lead to (1.1) being estimable by regression, but (1.2) was not, as z_{1t} was still correlated with u_{2t} owing to the fact that the errors u_{1t} and u_{2t} were correlated. Wold (1949),

[1] Namely that the number of excluded exogenous variables is greater or equal to the number of included endogenous variables.

(1951) seems to have been the first person to propose that the u_{1t} and u_{2t} be made uncorrelated. When allied with $a^0_{12} = 0$ this assumption defined what he termed a "recursive system". In such systems regression could be used to estimate both (1.1) and (1.2). Despite Wold's strong advocacy it has to be said that recursive systems did not catch on. The reason was probably due to the fact that the data available for estimation at the time was largely measured at a yearly frequency, and so it was hard to believe in a recursive structure. It was not until quarterly and monthly data began to proliferate that recursive systems became popular, with Sims (1980) forcefully re-introducing the idea.

Initially macro-econometric models were quite small. Although Tinbergen's (1936) model had 24 equations the representative of the next generation treated in textbooks - the Klein- Goldberger model - had just 15. In the public policy sphere however, macro-economic models became very large, possibly because of developments in computer software and hardware. A query often raised when these models were used was whether a variable could be regarded as being determined outside of the system, e.g. the money stock was often taken as exogenous, but the fact that central banks found it difficult to set the level of money suggested that it was not reasonable to treat it as exogenous. Even items such as tax rates were often varied in response to macro-economic outcomes and so could not be regarded as being completely exogenous.

These qualms led to some questioning of the way in which the macro modeling exemplified by the very large scale models of the 1960s and 1970s proceeded. Some of these might be traced back to Tinbergen's early work, where he proposed structural equations that involved expectations, and then replaced the expectations with a combination of a few variables. The question is why one would just use a few variables? Surely, when forming expectations of variables, one would use all relevant variables. This concern became particularly striking when rational expectation ideas started to emerge, as the prescription there was for expectations to be a combination of all lagged variables entering into a model, with none of them being excluded, although it might be that the weights on some variables could be quite small, leading to them being effectively excluded. One implication of this was that, if the weights needed to form expectations were unknown, it was doubtful if one could expect that the models would be identified, as no variables could be excluded from the structural equations. Therefore structural identification needed to be achieved by some other method, and there the notion of a recursive system became important.

The history just cited led to the principles of retaining lags of all variables in each structural equation but excluding some endogenous variables through the assumption of a recursive system. Formally there was no longer any distinction between endogenous and exogenous variables. To summarize the data all variables were taken to depend on the lags of all variables. Such a system had been studied by Quenouille (1957) and became known as a Vector Autoregression (VAR). As data was summarized by a reduced form, the VAR became that and its corresponding structural form was the Structural Vector Autoregression (SVAR).

After Sims (1980) SVARs became a very popular method of macroeconomic data analysis. Part of the appeal was their focus upon dynamic responses. The dynamic multipliers that were a feature of the older macro-econometric models were now re-named impulse responses, and the exogenous variables became the uncorrelated shocks in the SVAR. Because impulse responses were becoming very popular in theoretical macro economic work this seemed to provide a nice unification of theory and practice. Shocks rather than errors became the dominant perspective.

This monograph begins where the history above terminates. Chapter 2 begins by describing how to summarize the data from a VAR perspective and how to estimate this multivariate model. Many questions arise from matching this model structure to data, which can be loosely referred to as specification issues, e.g. choosing the order of the VAR, what variables should enter into it, should one restrict it in any way, and how might it need to be augmented by terms describing secular or specific events? Chapter 3 then turns to the usage of a VAR. Impulse response functions are introduced and forecasting is given some attention. Construction of these sometimes points to the need for the basic VAR structure outlined in Chapter 1 to be extended. Examples that fall into this category would be non-linear structures such as Threshold VARs, latent variables and time varying VARs. Each of these is given a short treatment in Chapter 3. To the extent to which the topics cannot be implemented in EViews 9 they are given a rather cursory treatment and reference is just made to where computer software might be found to implement them.

Chapter 4 starts the examination of structural VARs (SVARs). Basically, one begins with a set of uncorrelated shocks, since this is a key feature of modern structural models, and then asks how "names" can be given to the shocks. A variety of parametric restrictions are imposed to achieve this objective. Many issues are discussed in this context, including how to deal with stocks and flows, exogeneity of some variables, and the incorporation of "big data" features in terms of factors. Applications are given from the literature and a link is made with another major approach to macroeconometric modeling, namely that of Dynamic Stochastic General Equilibrium (DSGE) models.

In Chapter 5 the restriction that variables are stationary - which was used in Chapters 2-4 - is retained, but now parametric restrictions are replaced by sign restrictions as a way of differentiating between shocks. Two methods are given for implementing sign restrictions and computed by using both a known model of demand and supply and the small empirical macro model that was featured in earlier chapters. A range of problems that can arise with sign restrictions are detailed. In some cases a solution exists, in others the issue remains.

Chapter 6 moves on to the case where there are variables that are non-stationary, explicitly I(1), in the data set. This means that there are now permanent shocks in the system. However there can also be transitory shocks, particularly when I(0) variables are present. The combination of I(1) and I(0) variables also modifies the analysis, as it is now necessary to decide whether the structural shocks in the equation describing the I(0) variable is permanent or transitory. Examples are given of how to deal with the possibility that they are permanent both in the context of parametric and sign restrictions. Lastly, Chapter 7 takes up many of the same issues dealt with in the preceding chapter but with cointegration present between the I(1) variables. Now the summative model can no longer be a VAR but must be a Vector Error Correction Model (VECM) and there is a corresponding structural VECM (SVECM). To deal with the modeling issues it is useful to transform the information contained in the SVECM into an SVAR which involves variables that are changes in the I(1) variables and the error correction terms. Cointegration then implies some restrictions upon this SVAR, and these deliver instruments for the estimation of the equations. Two examples are taken from the literature to show how the methods work.

Chapter 2

Vector Autoregressions: Basic Structure

2.1 Basic Structure

Models used in macroeconomics serve two purposes. One is to be *summative,* i.e. to summarize the data in some coherent way. The other is to be *interpretative,* i.e. to provide a way to structure and interpret the data. In the early history of times series it was noticed that the outcome of a series at time t depended upon its past outcomes, i.e. the series was dependent on the past. It was therefore proposed that a simple model to capture this dependence would be the linear autoregression (AR) of order p

$$z_t = b_1 z_{t-1} + ... + b_p z_{t-p} + e_t,$$

where e_t was some shock (or what was then described as an "error") with a zero mean, variance σ^2 and which was not predictable from the past of z_t, i.e. all the dependence in z_t came from the lagged values of z_t. Consequently, it was natural that, when dealing with more than one series, this idea would be generalized by allowing for a system of autoregressions. One of the first to deal with this was Quenouille (1957) who investigated the Vector Autoregression of order p (VAR(p)):

$$z_t \quad = \quad B_1 z_{t-1} + .. + B_p z_{t-p} + e_t, \tag{2.1}$$

where now z_t and e_t are $n \times 1$ vectors and B_j are $n \times n$ matrices. Equation (2.1) specifies that any series depends on the past history of all the n series through their lagged values. When $p = 2$ there is a VAR(2) process of the form

$$z_t \quad = \quad B_1 z_{t-1} + B_2 z_{t-2} + e_t.$$

Letting $n = 2$ and expanding this out delivers the structure

$$\begin{aligned}\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} &= \begin{pmatrix} b^1_{11} & b^1_{12} \\ b^1_{21} & b^1_{22} \end{pmatrix}\begin{pmatrix} z_{1t-1} \\ z_{2t-1} \end{pmatrix} + \\ & \begin{pmatrix} b^2_{11} & b^2_{12} \\ b^2_{21} & b^2_{22} \end{pmatrix}\begin{pmatrix} z_{1t-2} \\ z_{2t-2} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix},\end{aligned}$$

where the superscript indicates the lag number and the subscripts refer to the equation and variable numbers. The assumptions made about the shocks e_t allow for them to be correlated:

$$\begin{aligned} E(e_{1t}) &= 0,\ E(e_{2t}) = 0 \\ var(e^2_{1t}) &= \sigma_{11},\ var(e^2_{2t}) = \sigma_{22} \\ cov(e_{1t}e_{2t}) &= \sigma_{12}. \end{aligned}$$

We will often work with either a $VAR(1)$ or a $VAR(2)$ in order to illustrate ideas, as nothing much is gained in terms of understanding by looking at higher values of p.

2.1.1 Maximum Likelihood Estimation of Basic VARs

The equations in (2.1) are often estimated using the Maximum Likelihood method. To find the likelihood it is necessary to derive the joint density of the variables $z_1, ..., z_T$. This will be termed $f(z_1, ..., z_T; \theta)$, showing its dependence on some parameters θ. Letting Z_{t-1} contain $e_1, ..., e_{t-1}$, the joint density can then be expressed as

$$f(z_1, ..., z_T; \theta) = f_0(Z_p; \theta) \prod_{t=p+1}^{T} f(z_t|Z_{t-1}; \theta),$$

where $f_0(Z_p)$ is the unconditional density of $z_1 ...\ z_p$ and $f(z_t|Z_{t-1}; \theta)$ is the density of z_t conditional on the past Z_{t-1}. Therefore, the log likelihood will be

$$L(\theta) = \ln(f(z_1, ..., z_T; \theta)) = \ln(f_0(Z_p; \theta)) + \sum_{t=p+1}^{T} \ln(f(z_t|Z_{t-1}; \theta)).$$

Because the second term increases with the sample size it might be expected to dominate the first, and so one normally sees it treated as an approximation to the log likelihood. Finally to give this a specific form some distributional assumption is needed for e_t. Making the density of e_t conditional upon Z_{t-1} be multivariate normal $N(0, \Omega_R)$ means that $f(z_t|Z_{t-1}; \theta) = N(B_1 z_{t-1} + .. + B_p z_{t-p}, \Omega_R)$, and so the approximate log likelihood will be

$$\begin{aligned} L(\theta) = cnst - \frac{T-p}{2} \ln|\Omega_R| - \\ \frac{1}{2} \sum_{t=p+1}^{T} (z_t - B_1 z_{t-1} - .. - B_p z_{t-p})' \Omega_R^{-1} (z_t - B_1 z_{t-1} - .. - B_p z_{t-p}), \end{aligned}$$

where *cnst* is a constant term that does not depend on θ.

Provided there are no restrictions upon B_j or Ω_R, each equation has exactly the same set of regressors, meaning that the approximate MLE estimates can be found by applying OLS to each equation in turn.[1] This makes the estimation of the basic VAR quite simple. If there are some restrictions then OLS would still provide an estimator but it would no longer be efficient. In that case one needs to maximize L to get the efficient estimator. We note that because the p initial conditions in $\ln(f_0(Z_p;\theta))$ have been ignored, the first p observations in the sample are discarded when OLS is used.

2.1.2 A Small Macro Model Example

An example we will use a number of times in this monograph involves a small macro model that has three variables - a GDP gap y_t (log GDP after linear detrending), inflation in the GDP deflator (π_t), and the Federal Funds rate (i_t). Data on these variables was taken from Cho and Moreno (2006) and runs from 1981/1-2000/1. An EViews workfile *chomoreno.wf1* contains this data and in it the variables are given the names *gap*, *infl* and *ff*. A VAR(2) fitted to this data would have the form

$$\begin{aligned} y_t = {} & b^1_{11}y_{t-1} + b^1_{12}\pi_{t-1} + b^1_{13}i_{t-1} + b^2_{11}y_{t-2} + \\ & b^2_{12}\pi_{t-2} + b^2_{13}i_{t-2} + e_{1t} \end{aligned} \tag{2.2}$$

$$\begin{aligned} \pi_t = {} & b^1_{21}y_{t-1} + b^1_{22}\pi_{t-1} + b^1_{23}i_{t-1} + b^2_{21}y_{t-2} + \\ & b^2_{22}\pi_{t-2} + b^2_{23}i_{t-2} + e_{2t} \end{aligned} \tag{2.3}$$

$$\begin{aligned} i_t = {} & b^1_{31}y_{t-1} + b^1_{32}\pi_{t-1} + b^1_{33}i_{t-1} + b^2_{31}y_{t-2} + \\ & b^2_{32}\pi_{t-2} + b^2_{33}i_{t-2} + e_{3t} \end{aligned} \tag{2.4}$$

The following screen shots outline the basic steps in EViews required to fit a VAR(2) to the Cho and Moreno data set. In this and following explanations of EViews procedures, **bold** means the command is available from the menu tab and is to be selected and clicked on using a pointer. Thus the first procedure involves opening the data set and then clicking on the sequence of commands: ***File*** $\rightarrow$ ***Open*** $\rightarrow$ ***EViews Workfile (Ctrl+O).*** The screen shot in Figure 2.1 below shows the resulting drop down menu.

Now locate the EViews data file *chomoreno.wf1* and click on it to open it in EViews. The result is shown in Figure 2.2. To fit the VAR(2) as in (2.2) - (2.4) issue the commands ***Quick*** $\rightarrow$***Estimate VAR*** and fill in the boxes as described in Figure 2.3. Note that a second order VAR requires one to state the range of lags that are to be fitted, i.e. 1 2 is entered into the lag intervals box. For a VAR(1) this would be stated as 1 1. Clicking the "OK" button then produces the following results in Figure 2.4

[1] This also shows that weaker conditions than normality of e_t can be assumed.

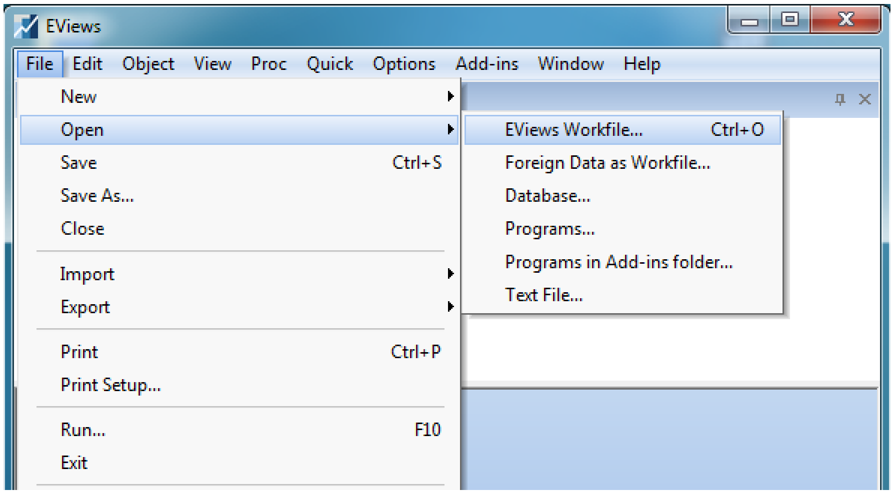


Figure 2.1: Opening A Workfile in EViews 9

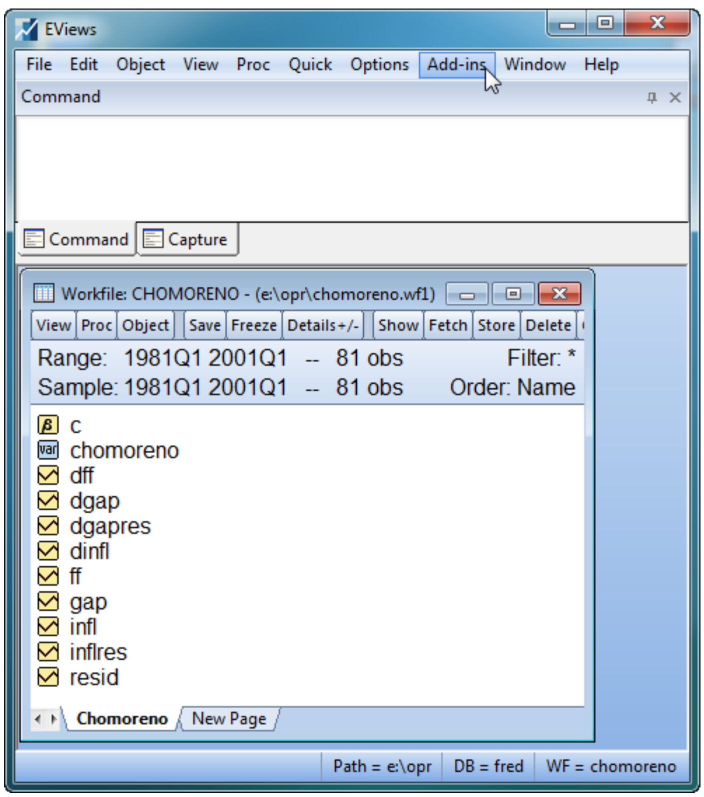


Figure 2.2: Chomoreno Data Set in EViews 9

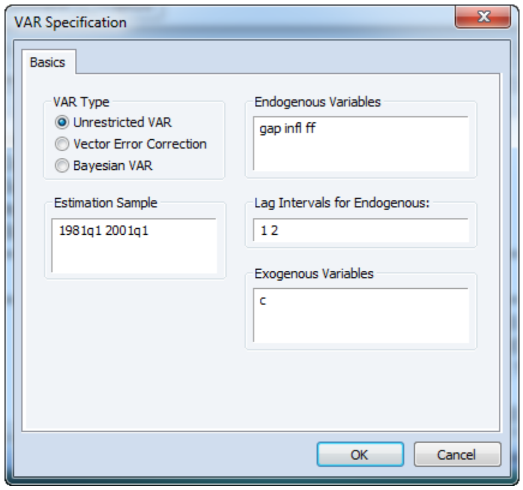


Figure 2.3: Specification of the Small Macro Model (VAR) in EViews 9

EViews also has a *command* line interface to estimate VAR models that is especially useful for replicating operations. After ensuring that the *chomoreno.wf1* file is open it is necessary to set the sample period for estimation. Since all the data is being used in this example, this can be easily done with the following EViews smpl command:

smpl @all

After that a VAR object called *chormoreno* can be created with the command

var chomoreno.ls 1 2 gap infl ff

In this case, estimation will be carried out using Ordinary Least Squares (OLS) (*i.e.* the *ls* command in EViews). The remaining items on the command ("1 2") refer to the order of the VAR (in this case 2) and the variables entering the VAR (i.e. gap infl and ff). Lastly, the contents of the chomoreno object (namely the estimation results) can be displayed using the show command:

show chomoreno

Typically, these commands will be placed in a program file (Figure 2.5) that can be executed by clicking on the ***Run*** tab button.

2.2 Specification of VARs

There are many issues that arise with VARs involving either some characteristics of the basic VAR outlined above or which represent extensions of it. The first two to be considered involve making a choice of

1. The order of the VAR.

2. The choice of variables to be included in the VAR, i.e. z_t.

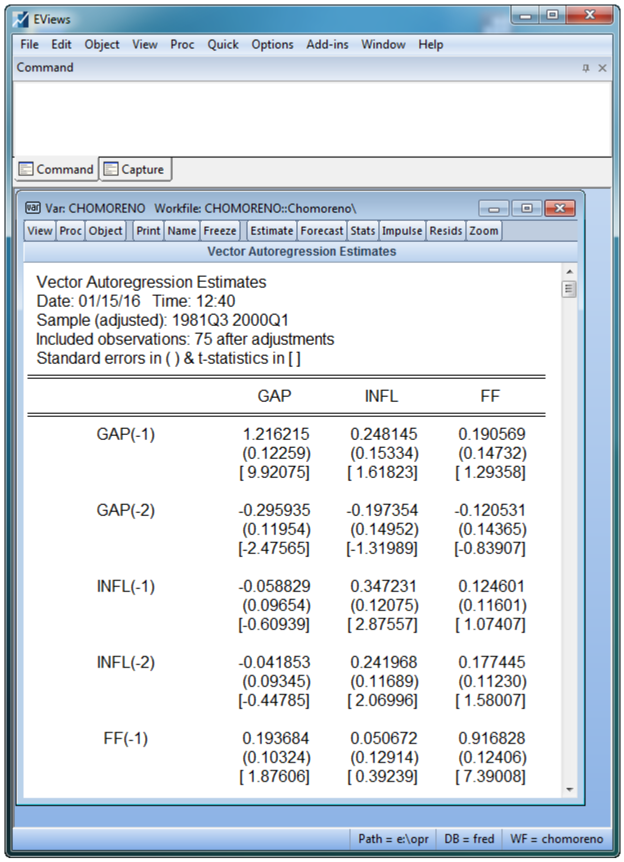


Figure 2.4: Results from Fitting a VAR(2) to the Small Macro Model Data Set

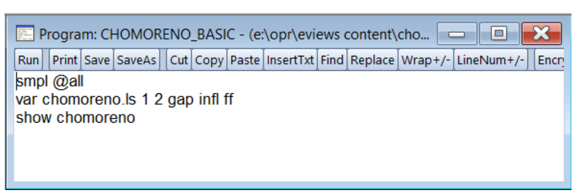


Figure 2.5: EViews Program to Replicate the Small Macro Model Results

In most instances a good deal of attention is paid to the first of these and little to the second, but it will be argued in Chapter 3 that the latter is at least as important. In fact, there is an interdependence between these two choices which has to be recognized.

2.2.1 Choosing p

There are basically three methods that have been employed to determine what p should be

1. By using some theoretical model.
2. By using a rule of thumb.
3. By using statistical criteria that trade off fit against the number of parameters fitted.

2.2.1.1 Theoretical models

Consider a small New Keynesian (NK) model for the determination of inflation (π_t), output (y_t) and the interest rate i_t in (2.5) - (2.7). This was the model that Cho and Moreno fitted.

$$y_t = \alpha_{yy}y_{t-1} + \beta_{yy}E_t(y_{t+1}) + \gamma_{yi}i_t + u_{yt} \tag{2.5}$$
$$\pi_t = \alpha_{\pi\pi}\pi_{t-1} + \beta_{\pi\pi}E_t(\pi_{t+1}) + \gamma_{\pi y}y_t + u_{pt} \tag{2.6}$$
$$i_t = \alpha_{ii}i_{t-1} + \gamma_{iy}y_t + \beta_{i\pi}E_t(\pi_{t+1}) + u_{it}. \tag{2.7}$$

In (2.5) - (2.7) if errors u_{yt}, u_{pt} and u_{it} are assumed to jointly follow a VAR(1) process then the solution for y_t, π_t and i_t will be a VAR(2). Indeed, it is the case that virtually all Dynamic Stochastic General Equilibrium (DSGE) models (of which the NK model is a representative) that have been constructed imply that the model variables jointly follow a VAR(2) - an exception being the model in Berg *et al.* (2006) which has a fourth order lag in the inflation equation. It is possible that the parameter values are such that the VAR(2) collapses to a VAR(1), but fairly rare. One instance where it does is when the errors have no serial correlation in them (and this was Cho and Moreno's assumption about them). Hence, if one has a DSGE model of an economy in mind, one would know what the potential set of variables to appear in a VAR would be, as well as the likely order of it. Generally, if data is quarterly a VAR(2) would probably suffice.

2.2.1.2 Rules of Thumb

Initially in practice one used to see people choosing $p = 4$ when working with quarterly data and $p = 6$ with monthly data. Provided n is small these are probably upper limits to the likely order. This is a consequence of the number of parameters that need to be estimated in each equation of the VAR - this being at least np (intercepts and other variables add to this count). Accordingly, a large p can rapidly become an issue unless there is a large sample size. It is probably unwise to fit more than $T/3$ parameters in each equation. Even when there is no relation between z_t and its lags the R^2 from applying OLS to each equation will be around the ratio of the number of parameters to the sample size. Indeed, in the limit when $np = T$, it is unity, regardless of the relationship between z_t and its lags. So the effective constraint is something like $np < \frac{T}{3}$ i.e. $3np < T$. With quarterly data we would often have no more than 100 observations, so putting $p = 4$ would mean that n cannot be chosen to be greater than 7.

2.2.1.3 Statistical Criteria

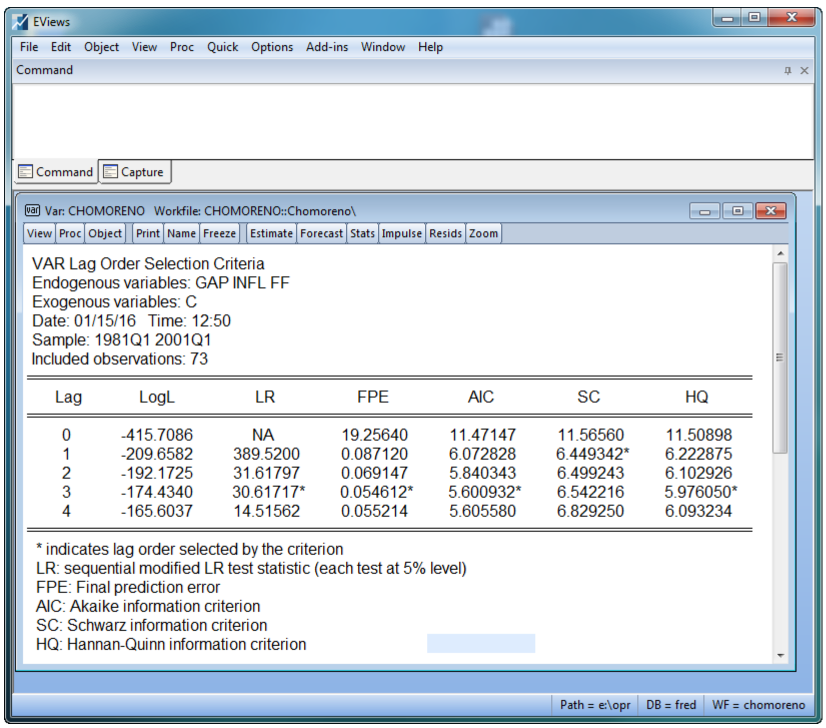
One might seek to choose p by seeing how well the data is fitted by the VAR(p) versus a VAR(q), where $p \neq q$. The problem is that one can get an exact fit by setting either p or q to T/n (T being the sample size). For this reason one wants to devise criteria that trade off fit and the number of parameters. There are a variety of such criteria and three are given in EViews - the Akaike Information Criterion (AIC), the Schwartz Bayesian Information Criterion (SC) and the Hannan-Quinn (HQ) Criterion. Using the log likelihood (L) as a measure of fit these criteria have the forms (where K is the number of parameters estimated),

$$\begin{aligned} AIC &: -2(\frac{L}{T}) + 2\frac{K}{T} \\ SC &: -2(\frac{L}{T}) + \frac{\ln(T)K}{T} \\ HQ &: -2(\frac{L}{T}) + 2\frac{\ln(\ln(T))K}{T}. \end{aligned}$$

If these criteria were applied to whether one should add extra regressors to a regression model, the rules would retain the regressors if the F statistic exceeded $(T - K - 1)(e^{2/T} - 1)$ (AIC) or $(T - K - 1)(e^{(\ln T)/T} - 1)$ (SC). Because $\frac{(e^{2/T}-1)}{(e^{(\ln T/T)}-1)} < 1$ this implies that AIC prefers larger models to SC. Consequently, we tend to prefer SC, as it seems unwise to estimate a large number of parameters with limited data. Notice that, because of the negative sign on the L term, we are trying to *minimize* each of the criteria.

To compute these in EViews for a range of values of p, after estimation of the VAR for a given order, click on ***View→Lag Structure→Lag Length Criteria.*** EViews then asks the user for “lags to include” which is designed so as to prescribe an upper limit for p. EViews then tests VAR orders up to that maximum order. The output using $p = 4$ is shown in Figure 2.6

In this screen the asterisk shows the minimum value for each criterion. Accordingly, for the small macro model the output from both HQ and AIC would point to a VAR(3), whereas SC flags a VAR(1). As said previously our preference is to select the most parsimonious model and that would be a VAR(1). However, as seen later with this data set, one might want to check if the VAR(1)



VAR Lag Order Selection Criteria
Endogenous variables: GAP INFL FF
Exogenous variables: C
Date: 01/15/16 Time: 12:50
Sample: 1981Q1 2001Q1
Included observations: 73

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-415.7086	NA	19.25640	11.47147	11.56560	11.50898
1	-209.6582	389.5200	0.087120	6.072828	6.449342*	6.222875
2	-192.1725	31.61797	0.069147	5.840343	6.499243	6.102926
3	-174.4340	30.61717*	0.054612*	5.600932*	6.542216	5.976050*
4	-165.6037	14.51562	0.055214	5.605580	6.829250	6.093234

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Figure 2.6: Choice of VAR Lag Length for the Small Macro Model

could be augmented in some way. Owing to serial correlation in the VAR errors, it might be that some extra lags need to be added on to some of the equations, even if not all of them.

2.2.2 Choice of Variables

There are two ways that this has been done

1. By institutional knowledge.
2. From theoretical models.

2.2.2.1 Institutional Knowledge

Often people in institutions doing macroeconomic modeling develop intuition over what variables are needed to adequately model the system. Early macro modelers argued that some monetary stock variable such as $M1$ would need to be added to the three variables in the small macro model discussed earlier to capture interactions in a closed economy. For a small open economy, it would be hard not to have the real exchange rate and foreign output in the list of variables appearing in the VAR. It also needs to be recognized that, in an open economy, there are independent measures of demand and supply, with the current account reflecting any discrepancy between them. Thus, as well as including a variable such as GDP in the VAR, it can be useful to add in a variable like Gross National Expenditure (GNE), with the latter playing the role of the "absorption" variable that appears in theoretical models of open economies.

2.2.2.2 Theoretical models

Just as happened with the selection of p, theoretical models can suggest what variables might appear in the VAR. Thus, from the New Keynesian perspective, one would choose y_t, π_t and i_t. One difficulty that often arises however is that such theoretical models often incorporate variables that are not easily measured, e.g. any DSGE model with a production side generally has in it the (unobserved) level of technology, and so this would appear among the z_t. Since one needs to be able to measure variables to put them in standard VAR packages it is tempting to just select a sub-set of the model variables. But this can have consequences, and we will look at these later in Chapter 3.

2.2.3 Restricted VAR's

One reason to be careful when choosing a high value of p is that it might either be a reflection of an n that is too small or that the wrong variables have been chosen. A further complicating factor is that it also may be that the higher lags belong in some of the VAR equations and not others i.e. some of the elements of the b_j are zero. We will refer to this as a restricted VAR. Such exclusion restrictions can be for theoretical reasons. For example take the NK model in (2.5) - (2.7) with $\alpha_{yy} = 0, \alpha_{\pi\pi} = 0$ (not an uncommon choice for those estimating this model). This restriction arises from the fact that, if there is no serial correlation in the shocks of the NK model, then the solution is a VAR(1) with the coefficients on y_{t-1} and π_{t-1} being zero in all equations.

2.2.3.1 Setting Some Lag Coefficients to Zero

Statistical evidence might also show that some of the coefficients in B_j should be set to zero. To investigate that possibility one needs to examine the individual equations. Looking at these in the context of the equations in the three variable VAR(2) fitted to the Cho-Moreno data, we find that the t ratios for y_{t-2} and i_{t-2} in the y_t equation are -2.53 and 1.98 respectively, while in the other two equations second lags of variables are insignificant. Now, as seen earlier, a VAR(1) would have been selected with the SC criterion, while the results just mentioned suggest that such a VAR might need to be extended in order to have second lags of variables in some but not all of the equations. Such VARs could be termed *restricted* and essentially constitute an *unbalanced* (in the lag variables entering into equations) *VAR* structure. In relation to restricted VARs the EViews manual comments that using the option *Make System* in the **Procs** menu means it would be possible to account for the restrictions on the system in estimation. But, if you follow this route, you are moved out of the VAR object into the SYSTEM object, and so none of the VAR options re computing impulse responses etc. are immediately available. Nevertheless it is possible to handle restricted VARs and to compute impulse responses using a somewhat cumbersome procedure and the addition of a special program written in the EViews language. We will look at this in the next sub-section.

It is not entirely clear what the use of restricted VARs is. One instance may be if the VAR is to be used for forecasting, since the large number of parameters in VARs makes for rather imprecise forecasts, and retaining only those variables that have a significant role is likely to be important. However, the Bayesian approach to this which generates BVARs that give low rather than zero weight to such regressors seems to have much better forecasting properties.[2]

EViews can test for the exclusion of all the lagged values of a single variable from a VAR. To do this, after estimation use ***View*$\rightarrow$ *Lag Structure*$\rightarrow$ *Lag Exclusion Tests*.** Unfortunately, this is simply a test. If one wants to impose the restriction that a specific lag is absent then it is necessary to move to the SYSTEM object, and so it will be necessary to later ask how one proceeds in such a case. It should be noted that there is a literature which suggests that unbalanced VARs could be selected with algorithms that automate the choices. The best-known of these is PC-Gets - see Hendry and Krolzig (2005) - which has been applied to produce parsimonious VAR structures by deleting lags in variables if they fail to meet some statistical criteria. Heinlin and Krolzig (2011) give an application of this methodology to find a VAR to be used for examining over-shooting in exchange rates.

[2]We will deal with BVARs in the next chapter.

2.2.3.2 Imposing Exogeneity- the VARX Model

To date the traditional classification of variables as endogenous and exogenous has not been used. It is often said that all variables are treated as endogenous, although this is not strictly correct. When performing macroeconomic analysis there are clearly cases where variables are best thought of as exogenous, and that should determine how the VAR is formulated. Perhaps the clearest example of this would be in the context of a small open economy, where the foreign variables would be expected to affect the domestic ones but not conversely i.e. the foreign variables would be determined by their own lag values and not those of domestic variables. This clearly imposes zero restrictions upon the B_j in a VAR. Such a structure leads to the *VARX model* (the X meaning it is a VAR with exogenous variables), where one (or more) variables are treated as being exogenous relative to another set of variables.

To be more concrete about this consider a data set for Brazil called *brazil.wf1*, which has quarterly macroeconomic data spanning 1999:2-2008:4 (the era covering the introduction of an inflation target).[3] Here y_t is an output gap, n_t is an absorption gap, π_t (called $infl_t$ in the data set) is an inflation rate adjusted for the Brazilian central bank's target inflation, i_t(called int_t in the data set) is an interest rate adjusted in the same way, rer_t is a real exchange rate, $ystar$ is a foreign output gap and rus_t is a real foreign short-term interest rate. The data set was used in Catao and Pagan (2011). The SVAR is in terms of y_t, π_t, i_t, and rer_t, with $ystar_t$ and rus_t being treated as exogenous. Because the sample size is small we use a VAR(1).

To see what the resulting VARX system looks like consider the domestic output gap equation. It has the form

$$y_t = a_{11}^1 y_{t-1} + a_{12}^1 n_{t-1} + a_{13}^1 \pi_{t-1} + a_{14}^1 i_{t-1} + a_{15}^1 rer_{t-1} + \gamma_{11}^0 y_t^* + \gamma_{12}^0 rus_t + e_{1t}, \tag{2.8}$$

where γ_{ij}^0 shows the contemporaneous impact of the $j'th$ exogenous variable upon the $i'th$ endogenous variable. The screen shot in Figure 2.7 shows how this is implemented using the Brazilian data in *brazil.wf1*

It may be necessary to construct data on lags of some of the exogenous variables, e.g. it might be that y_{t-1}^* should be in the VARX equations along with y_t^* and, in that case, one would need to enter *ystar*, *ystar(-1)* in the "exogenous variables" box when defining the model. Notice that the exogenous variable has to appear *in every equation* for the endogenous variables. Using standard pull-down menus one cannot have them in one equation but not another, so we need to describe how this constraint can be relaxed as it is the basis for handling any restricted VARs.

[3] We will use 1999:2 to mean the second quarter of 1999. Later with monthly data 1999:2 will mean the third month.

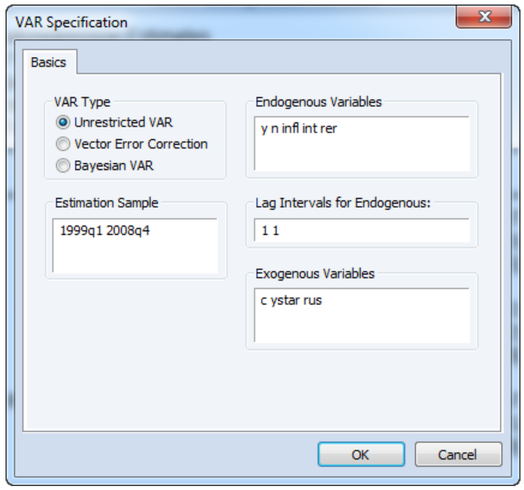


Figure 2.7: Specification of a VARX Model for the Brazilian Macro Data Set

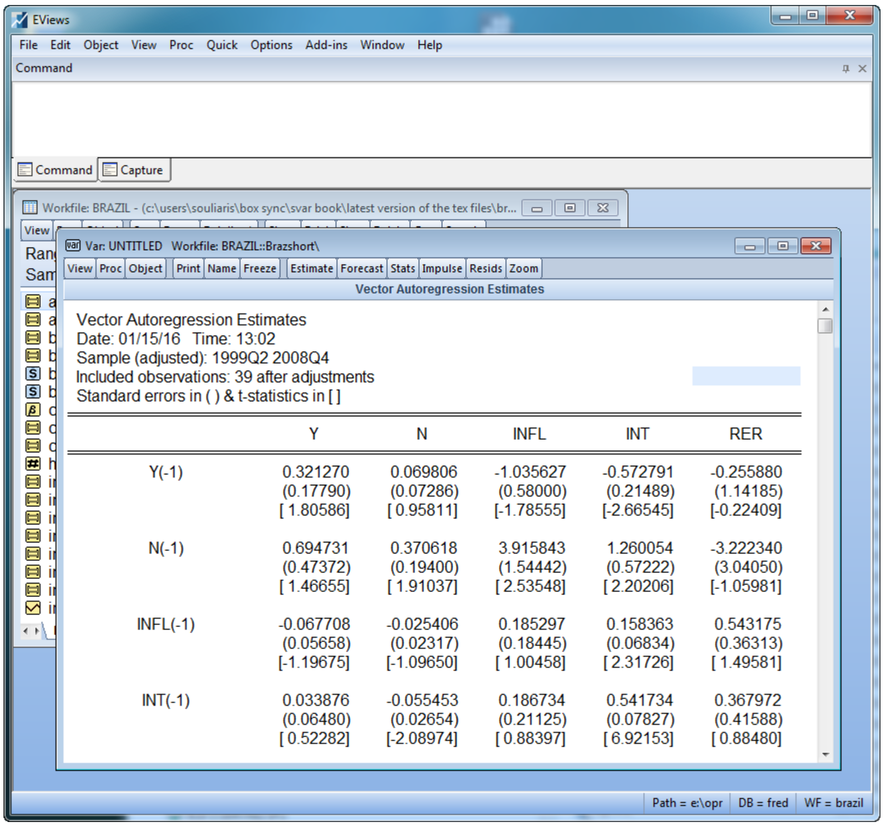
Suppose that we want the foreign interest rate rus_t to only appear in the real exchange rate equation. Running the VAR in Figure 2.7 above we will get the results in Figure 2.8. Subsequently, choosing the ***Proc***$\rightarrow$***Make System***$\rightarrow$ ***Order by Variable*** will push the resulting VAR specification to the EViews system estimator (Figure 2.9).

Now this system needs to be edited so as to produce a VAR that has rus absent from all equations except that for rer. This means we need to re-number the coefficients to be estimated. The resulting system is shown in Figure 2.10.

Then choosing ***Estimate*** $\rightarrow$ ***Ordinary Least Squares*** will give the results in Figure 2.11.[4]

These estimates now need to be mapped into the VAR matrices B_1 and F in the system $z_t = B_1 z_{t-1} + F\xi_t$, where $\xi_t = \begin{bmatrix} e_t \\ z_t^* \end{bmatrix}$, z_t^* are the exogenous variables (omitting deterministic ones like dummy variables, trends and constants) and F is a matrix. The impulse responses computed directly from the system will be to one unit changes in e_t and z_t^*, but these can be made one standard deviation changes by adjusting the elements in F appropriately. Thus, in our example, where there are five shocks e_t and two exogenous variables y_t^* and rus_t, for the first equation the (1, 6) element in F would be $C(7)$. However if a change equal to one standard deviation of y_t^* was desired it would be necessary to set it to

[4]One could also choose **Full Information Maximum Likelihood.** Because the foreign interest rate is excluded from some of the VAR equations OLS and FIML will no longer be the same. FIML is a more efficient estimator but OLS is consistent.



Vector Autoregression Estimates
Date: 01/15/16 Time: 13:02
Sample (adjusted): 1999Q2 2008Q4
Included observations: 39 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	Y	N	INFL	INT	RER
Y(-1)	0.321270	0.069806	-1.035627	-0.572791	-0.255880
	(0.17790)	(0.07286)	(0.58000)	(0.21489)	(1.14185)
	[1.80586]	[0.95811]	[-1.78555]	[-2.66545]	[-0.22409]
N(-1)	0.694731	0.370618	3.915843	1.260054	-3.222340
	(0.47372)	(0.19400)	(1.54442)	(0.57222)	(3.04050)
	[1.46655]	[1.91037]	[2.53548]	[2.20206]	[-1.05981]
INFL(-1)	-0.067708	-0.025406	0.185297	0.158363	0.543175
	(0.05658)	(0.02317)	(0.18445)	(0.06834)	(0.36313)
	[-1.19675]	[-1.09650]	[1.00458]	[2.31726]	[1.49581]
INT(-1)	0.033876	-0.055453	0.186734	0.541734	0.367972
	(0.06480)	(0.02654)	(0.21125)	(0.07827)	(0.41588)
	[0.52282]	[-2.08974]	[0.88397]	[6.92153]	[0.88480]

Figure 2.8: Brazilian VAR Results, 1992:2 2008:4

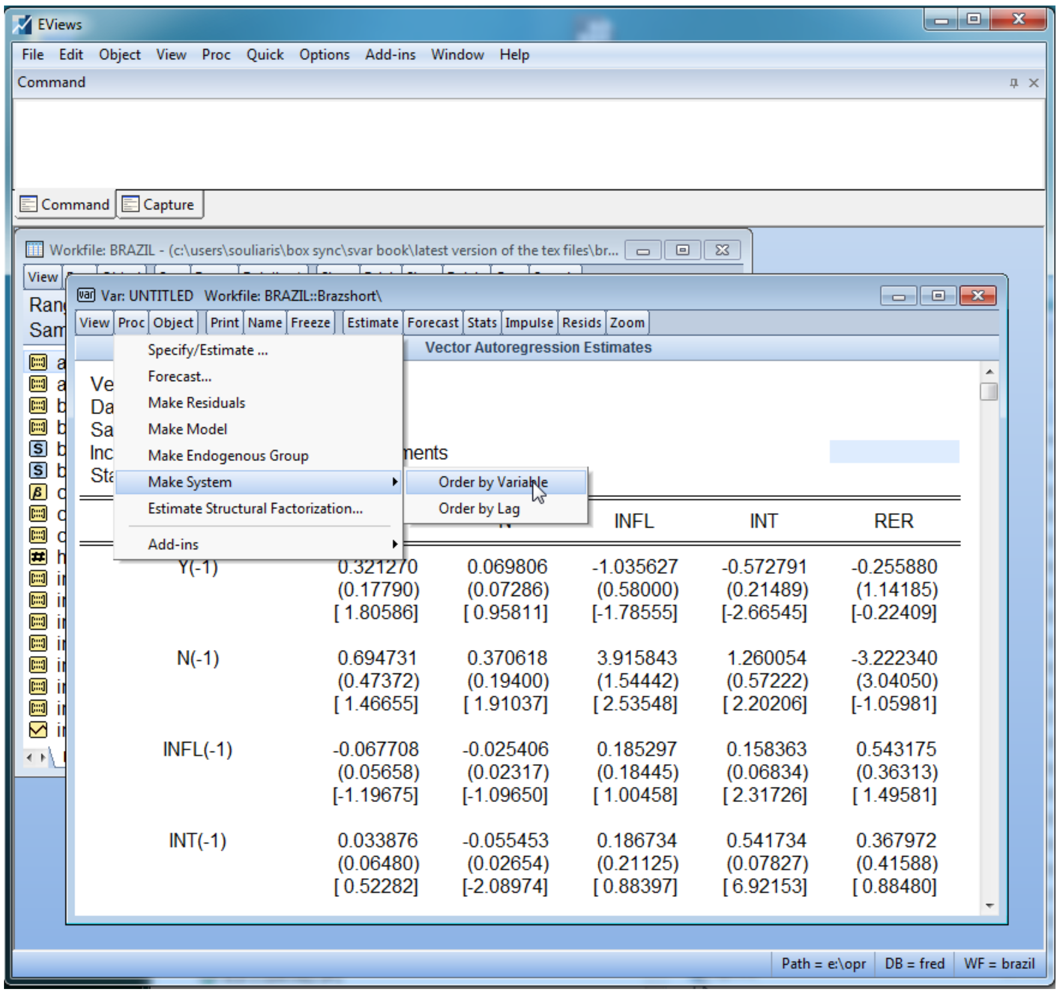


Figure 2.9: Moving to the System Estimator in EViews

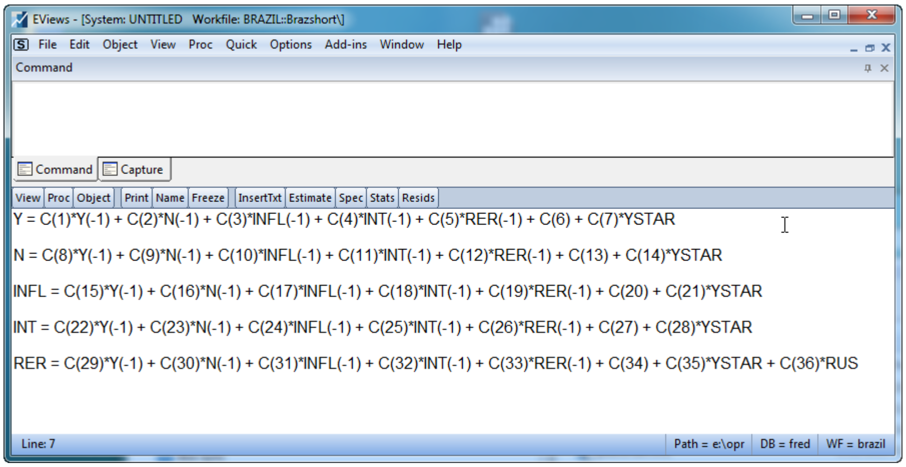


Figure 2.10: System Representation of the Brazilian VAR with Exogeneity Restrictions

$C(7) * std(y_t^*)$. The program *restvar.prg* shows how C is mapped into B_1 and F and how impulse responses are then computed.

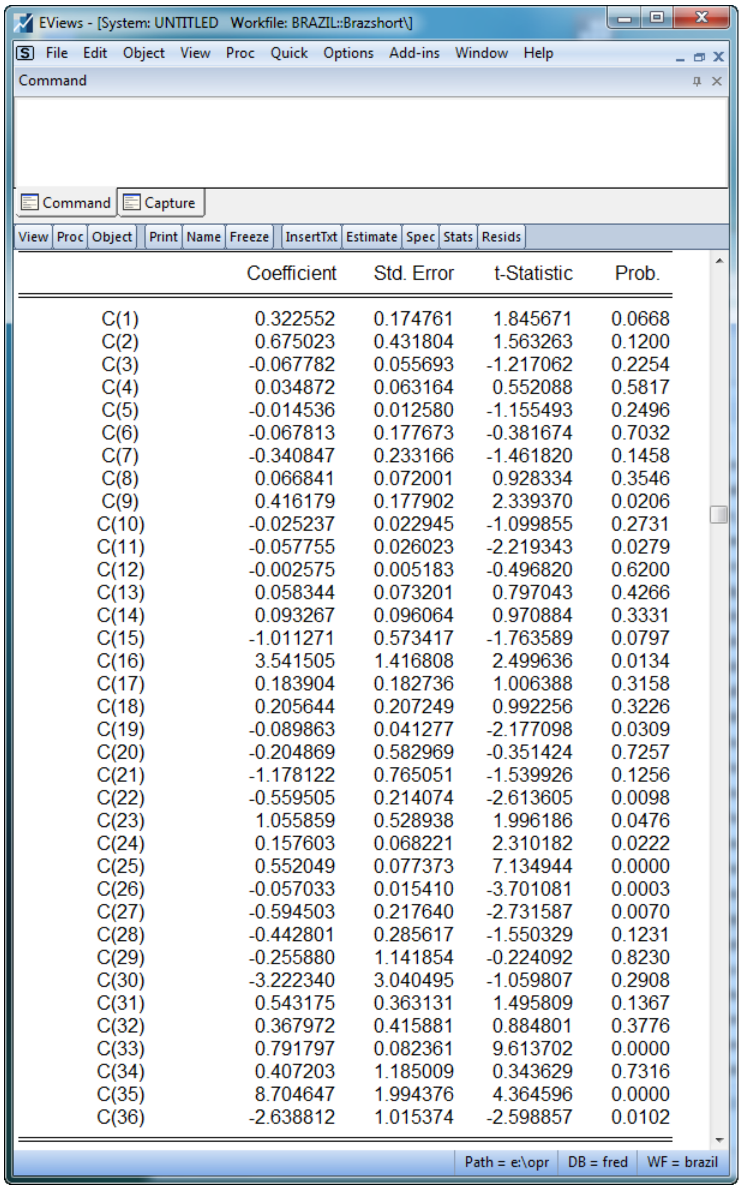
Our final comment on the VARX model is that exogenous variables are effectively being classified in that way because there is no equation for them in the VAR. An example of this would be Iacoviello (2005) who has a VAR in four variables - a GDP gap, rate of change of the GDP deflator, detrended real house prices and the Federal Funds rate. He then adds a variable - the log of the Commodity Research Bureau (CRB) commodity spot price index - on to the system. It might be argued that this would be endogenous for the US economy, but the analysis he performs is conditional upon commodity prices, so it is the VARX model that is relevant.

2.2.4 Augmented VARs

2.2.4.1 Trends and Dummy Variables

In VAR analysis the concern is to capture the joint behavior of a set of variables. Sometimes there can be specific events that cause movements in some of the variables but not others. Examples would be effects due to demographic features, running events like the Olympic Games, wars, and the introduction of new tax rates or special levies. Although these events can be thought of as exogenous it might be difficult to construct a specific series for them, leading to their being handled by the introduction of trend and dummy variables.

One problem is how they should be introduced into VARs. Including them as exogenous variables in the VAR specification means that they will appear in every equation of the VAR, and often this is not sensible. A good example of this is the Iacoviello (2005) study mentioned above who added to his four



	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.322552	0.174761	1.845671	0.0668
C(2)	0.675023	0.431804	1.563263	0.1200
C(3)	-0.067782	0.055693	-1.217062	0.2254
C(4)	0.034872	0.063164	0.552088	0.5817
C(5)	-0.014536	0.012580	-1.155493	0.2496
C(6)	-0.067813	0.177673	-0.381674	0.7032
C(7)	-0.340847	0.233166	-1.461820	0.1458
C(8)	0.066841	0.072001	0.928334	0.3546
C(9)	0.416179	0.177902	2.339370	0.0206
C(10)	-0.025237	0.022945	-1.099855	0.2731
C(11)	-0.057755	0.026023	-2.219343	0.0279
C(12)	-0.002575	0.005183	-0.496820	0.6200
C(13)	0.058344	0.073201	0.797043	0.4266
C(14)	0.093267	0.096064	0.970884	0.3331
C(15)	-1.011271	0.573417	-1.763589	0.0797
C(16)	3.541505	1.416808	2.499636	0.0134
C(17)	0.183904	0.182736	1.006388	0.3158
C(18)	0.205644	0.207249	0.992256	0.3226
C(19)	-0.089863	0.041277	-2.177098	0.0309
C(20)	-0.204869	0.582969	-0.351424	0.7257
C(21)	-1.178122	0.765051	-1.539926	0.1256
C(22)	-0.559505	0.214074	-2.613605	0.0098
C(23)	1.055859	0.528938	1.996186	0.0476
C(24)	0.157603	0.068221	2.310182	0.0222
C(25)	0.552049	0.077373	7.134944	0.0000
C(26)	-0.057033	0.015410	-3.701081	0.0003
C(27)	-0.594503	0.217640	-2.731587	0.0070
C(28)	-0.442801	0.285617	-1.550329	0.1231
C(29)	-0.255880	1.141854	-0.224092	0.8230
C(30)	-3.222340	3.040495	-1.059807	0.2908
C(31)	0.543175	0.363131	1.495809	0.1367
C(32)	0.367972	0.415881	0.884801	0.3776
C(33)	0.791797	0.082361	9.613702	0.0000
C(34)	0.407203	1.185009	0.343629	0.7316
C(35)	8.704647	1.994376	4.364596	0.0000
C(36)	-2.638812	1.015374	-2.598857	0.0102

Figure 2.11: Ordinary Least Squares Estimates of the Restricted VARX Model

variable VAR system not only commodity prices but also a time trend and a dummy variable that shifted the intercept from 1979Q4. The addition of the time trend looks odd given that both GDP and house prices have already been Band-Pass filtered, as this filter removes a linear trend. The major effect of adding the trend term to the VAR is upon the actual Federal Funds rate, as it is highly significant in that equation but no others. Consequently, one is effectively working with a 'detrended' Federal Funds rate and this seems questionable. Such an action certainly needs to be defended. This illustration raises the issue that it is important to give a rationale for the form of any variables entered into a VAR system, and one needs to recognize that the introduction of any form of exogenous variables can change the nature of the endogenous variables. Again, this illustrates the need for software that allows exogenous variables to affect only some of the VAR equations.

More generally, one might have changes in the joint relationship between series that need to be accounted for. These changes can be one of two types. First, it is possible that the parameters characterizing the unconditional densities of z_t (the unconditional moments) change. These changes could last for short or long periods and we will refer to them as *breaks*. The most important parameters to exhibit breaks would be the intercepts in the VAR equations, as that causes a break in the mean. Nevertheless, sometimes shifts in the variances of the errors occur for one of more variables, e.g. for GDP growth during the 20 or so years of the Great Moderation, and for interest rates during the Volcker experiment from 1979-1982. A second type of variation is what we will describe as *shifts* and these will be taken to occur when the parameters of the *conditional densities* change.

Looking first at breaks the problem is to locate where they happen. Once found we can generally handle them with dummy variables. Sometimes institutional events will tell us where to position the dummy variables, e.g. working with South African macro-economic data there were clear breaks post-Apartheid. In other cases one might ask whether the data can shed light on where the breaks occur. Some tests along these lines are available after a regression has been run. To access these click on the commands ***View→Stability Diagnostics*** and they will appear as in Figure 2.12. Of the tests that are made available the Chow break point test is good if one has some idea of where the breaks occur. If there is no hypothesis on this then one could look at either the Quandt-Andrews or the Bai-Perron Multiple break point test. The main disadvantage of the latter two is that the statistics are based on asymptotic theory, so that one needs a reasonable sample size. Moreover, the break points need to be away from the beginning and end of the sample since, for a given break point of T^*, one compares the estimated regression using data from $1...T^*$ with that for $T^*+1...T$. Both tests do insist that the user only search for breaks within the interior of the sample, i.e. T^* cannot be too close to the start of the sample.

Applying these tests to (say) the *gap* equation in the VAR, it is found that the last two tests suggest that a break happened around 1984Q4. This is probably too close to the beginning of the sample to decide that there was a break, as it would mean that one of the regressions fitted to test that the break point is

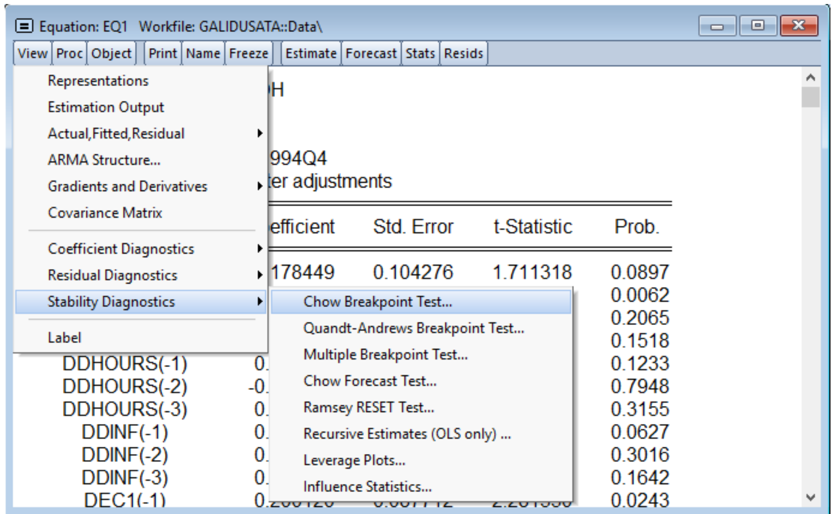


Figure 2.12: Screen Shot of the Options for Testing Breaks in the Coefficients of the VAR

there would only be using sixteen observations to fit seven parameters.

Sometimes it can pay to examine the recursive estimates of the regression parameters. In this case there seems to be a possibility of some mild drift. The *infl* equation shows no evidence of changes, but that for the Federal Funds rate does show some evidence, although again it is around 1985Q1. In general the evidence of any parameter change doesn't seem strong enough for any dummy variables to be added to the small macro VAR.

Lastly it can be useful to investigate breaks using programs that are not EViews. Hendry *et al.* (2008) have argued for the addition of many dummy variables (termed impulse indicators*)*, I_t, which take the values 1 at t and zero otherwise. They use a strategy whereby many of these are added to the VAR equations and then the equations are simplified using the algorithm in *Autometrics*. Often however this "impulse indicator saturation followed by simplification" approach leads to a lot of dummy variables for which there is no obvious rationale. It may be that they either reflect variations in the unconditional moments of the data or it may be that they are accounting for variations in the data that are not captured with the past history of the variables. In the latter instance they could be acting so as to make the shocks "better behaved", i.e. closer to being normally distributed. It is unclear whether one should just omit them and allow the shocks to be non-normal.

Shifts in the conditional densities are sometimes interpreted as coming from the fact that the densities depend on a set of *recurrent events,* e.g. it may be argued that the VAR coefficients differ either between recession and expansion periods. Another possibility is for the defining event to involve some threshold in an observable variable. Because these events recur with a certain probability it is generally the case that the unconditional densities of the z_t would have constant moments, i.e. there would be no breaks. A recurrent event like recessions would require augmentation of the VAR equations with terms involving an indicator that is unity during recessions and is zero otherwise (these can also be interacted with the lags). As we will see later one has to be careful when augmenting VARs with indicator variables that are constructed from endogenous variables in the VAR. Because recurrent events essentially induce non-linear structure into the VAR we will discuss them later under that heading.

2.2.4.2 With Factors

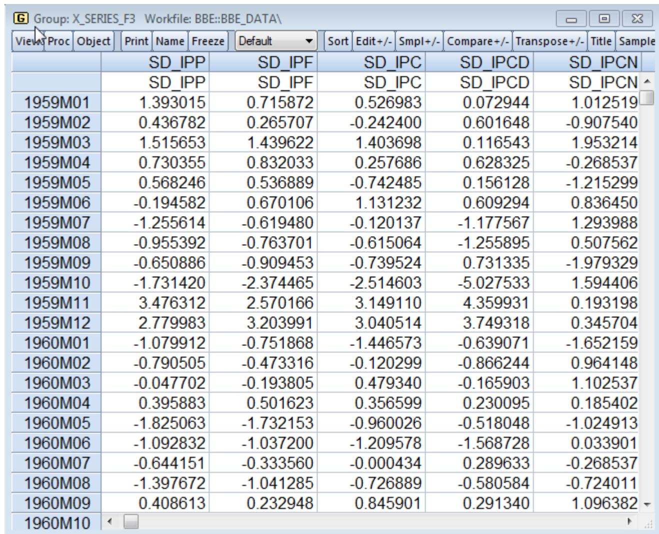
Often many variables are available to an analyst which are expected to influence the macro economy. Thus financial factors and confidence might be important to decisions. Because there is rarely a single measure of these there is a tendency to utilize many approximate measures, particularly involving data surveying the attitudes of financial officers, households or business men. There are far too many of these measures to put them all into a VAR, and so some prior aggregation needs to be performed. For a small system involving macroeconomic variables such as the unemployment rate, industrial production and employment growth, Sargent and Sims (1977) found that two dynamic factors could explain 80% or more of the variance of these variables. Bernanke *et al.* (2005) extended this approach.

In the Bernanke *et al.* variant one begins by assuming that there are m common factors F_t present in a set of N variables X_t. Then X_t has a factor structure and the factors are taken to follow a VAR. This VAR may include some *observable* variables y_t i.e. the system is

$$X_t = \Lambda F_t + v_t \tag{2.9}$$
$$F_t = B_{11}^1 F_{t-1} + B_{12}^1 y_{t-1} + \varepsilon_{1t} \tag{2.10}$$
$$y_t = B_{21}^1 F_{t-1} + B_{22}^1 y_{t-1} + \varepsilon_{2t}. \tag{2.11}$$

Clearly if F_t was available a VAR could be fitted to the data on F_t and y_t. Bernanke *et al.* proposed that one use the m principal components of X_t (PC_t) in place of F_t. Bai and Ng (2006) show that PC_t converges to F_t as $N \to \infty$. Generally N has to rise at a slower rate than the sample size T. Moreover, Bai and Ng show that the standard errors of the estimated coefficients of Λ and B_{ij} are not asymptotically affected by the fact that PC_t can be thought of as a generated regressor. This approach has become known as a Factor Augmented VAR (FAVAR).

EViews can compute the principal components of a set of data X_t and the commands to do this are now discussed. First, the data set used by Bernanke is *bbedata.wf1* and it needs to be opened. There are 120 variables in this data set, 119 in X_t and a single one (the Federal Funds rate) in y_t. They then standardized all the data i.e. mean corrected each of the series, followed by a scaling of those variables by their standard deviations. The resulting series are



	SD_IPP	SD_IPF	SD_IPC	SD_IPCD	SD_IPCN
	SD_IPP	SD_IPF	SD_IPC	SD_IPCD	SD_IPCN
1959M01	1.393015	0.715872	0.526983	0.072944	1.012519
1959M02	0.436782	0.265707	-0.242400	0.601648	-0.907540
1959M03	1.515653	1.439622	1.403698	0.116543	1.953214
1959M04	0.730355	0.832033	0.257686	0.628325	-0.268537
1959M05	0.568246	0.536889	-0.742485	0.156128	-1.215299
1959M06	-0.194582	0.670106	1.131232	0.609294	0.836450
1959M07	-1.255614	-0.619480	-0.120137	-1.177567	1.293988
1959M08	-0.955392	-0.763701	-0.615064	-1.255895	0.507562
1959M09	-0.650886	-0.909453	-0.739524	0.731335	-1.979329
1959M10	-1.731420	-2.374465	-2.514603	-5.027533	1.594406
1959M11	3.476312	2.570166	3.149110	4.359931	0.193198
1959M12	2.779983	3.203991	3.040514	3.749318	0.345704
1960M01	-1.079912	-0.751868	-1.446573	-0.639071	-1.652159
1960M02	-0.790505	-0.473316	-0.120299	-0.866244	0.964148
1960M03	-0.047702	-0.193805	0.479340	-0.165903	1.102537
1960M04	0.395883	0.501623	0.356599	0.230095	0.185402
1960M05	-1.825063	-1.732153	-0.960026	-0.518048	-1.024913
1960M06	-1.092832	-1.037200	-1.209578	-1.568728	0.033901
1960M07	-0.644151	-0.333560	-0.000434	0.289633	-0.268537
1960M08	-1.397672	-1.041285	-0.726889	-0.580584	-0.724011
1960M09	0.408613	0.232948	0.845901	0.291340	1.096382
1960M10					

Figure 2.13: Standardized Variables from the Bernanke *et al.* (2005) Data Set

given names like *sd_ip,* which is standardized growth in industrial production. Standardization is done using the *standard.prg* file and this can be implemented using the ***RUN*** command. The standardized series are then grouped into *x_series_f3* to represent the transformed X_t, and there is also one series *sd_fyff* to represent y_t. Figure 2.13 shows some of the standardized variables.

Now click ***Proc***$\rightarrow$ ***Make Principal Components***. The screen in Figure 2.14 appears and names must be given to the scores (principal components) and the loading matrix (Λ in (2.9)).

We compute three principal components from the total set of 119 variables (excluding the Federal Funds Rate which is y_t). It should be noted that a series such as the standardized growth in industrial production can be represented as a function of the three components. To do so we use the loadings for this variable giving $ip_t = .2126F^s_{1t} - .009F^s_{2t} + .1151F^s_{3t}$. Of course there are other factors (principal components) than the three computed but they are orthogonal to these three.

2.3 Conclusion

Chapter 2 has set out our basic summative model - the VAR. For this to be used data has to be stationary and so care needs to be taken when selecting the variables to enter the VAR as well as determining its maximum lag length. Of course it may be necessary to work with VARs that are restricted in some way,

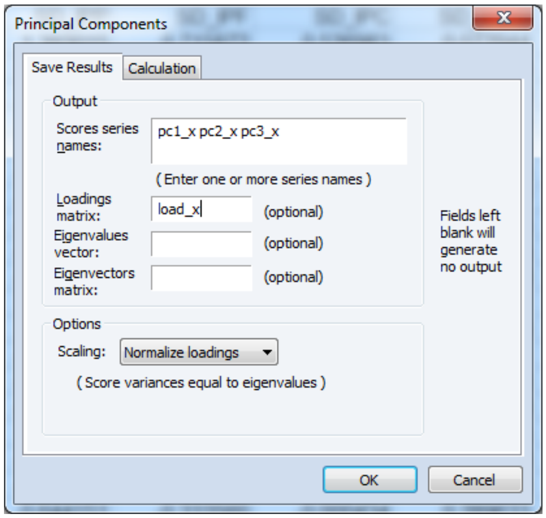


Figure 2.14: Extracting Three Principal Components from the Bernanke *et al.* (2005) Data Set

e.g. by having different maximum lags of variables in each equation. Handling restrictions like this can be done by moving out of the VAR object in EViews to the SYSTEM object, but it needs to be done with care. We will return to this strategy in later chapters. For the moment the presumption should be that unless there are very good reasons to impose restrictions, the VAR should be kept in its most general form so as to capture the underlying dynamics.

Chapter 3

Using and Generalizing a VAR

3.1 Introduction

There are many uses in the literature once a VAR for a set of variables and a given order has been fitted to data. One of these has been testing for Granger Causality and that is covered in the next section. After that attention is turned to using a VAR for forecasting in Section 3. Here a central problem becomes the number of parameters being fitted, and it has become customary to use Bayesian VARs that utilize Bayesian ideas to effectively constrain the values one might have for the B_j. Section 4 deals with that. Section 5 looks at the computation of impulse responses to the errors of the fitted VAR and how one is to describe the uncertainty in those values. After dealing with the use of VARs we turn to how one is to account for a variety of issues involving latent variables, non-linearities, and allowing for shifts in the conditional densities.

3.2 Testing Granger Causality

VARs have a number of uses. Often they are applied to testing for *Granger Causality,* i.e. whether one variable is useful for predicting another. Technically the question being posed is whether the past history (lags) of a variable y_{2t} influences y_{1t}. If it does then y_{2t} is said to cause y_{1t}. Mostly this is implemented as a bivariate test and involves regressing y_{1t} on lags of y_{1t} and y_{2t} followed by a test of whether the latter are zero. If this hypothesis is accepted there is no Granger causality. One could also introduce a third variable y_{3t} and ask whether the lags in y_{2t} and y_{3t} help explain y_{1t}, and that involves testing if the lag coefficients of both y_{2t} and y_{3t} are jointly zero. Granger Causality tests can therefore be implemented in EViews after fitting a VAR by clicking on the commands ***View→ Lag Structure→ Granger Causality/Block Exogeneity***

Tests. If there were three variables in the VAR (as above) the tests are given for deleting lags of y_{2t} and y_{3t} both separately and jointly. Clearly one is testing whether some of the elements in B_j are zero. Accordingly, if it is accepted that these parameters are zero, there would be a restricted VAR, and it was pointed out in the last chapter that EViews does not handle restricted VARs very well, at least when using its pull down menus.

To return to a theme of the last chapter suppose that y_{1t} were foreign variables and y_{2t} were domestic variables. Then the exogeneity of foreign variables for a small open economy could be thought of as foreign variables Granger causing the domestic variables, but not conversely. It does not seem very sensible to test for this foreign factor exogeneity if the economy is genuinely small, as the combination of limited data and the testing of a large number of parameters will most likely lead to rejection of the hypothesis that some of them are zero. Thus, for the Brazilian data of Chapter 2, if we test that the domestic variables have a zero effect on the foreign output gap the F test has a p value of .06. Given these qualms it is unclear what one learns from the many Granger Causality studies that have been performed. As Leamer (1984) pointed out many years ago what is being tested is whether one variable precedes another. As he says "Christmas cards precede Christmas but they don't cause Christmas". Of course Granger had a very specific definition of causality in mind, namely that one variable could improve the forecasts of another if it was used, and so his method of testing made sense in such a context.

3.3 Forecasting using a VAR

One of the most important uses of a VAR is producing forecasts. Assuming that the VAR fitted to the data is of order p, viz:

$$z_t = B_1 z_{t-1} + ... + B_p z_{t-p} + e_t, \;\; e_t \sim N(0, \sigma^2)$$

a one-period-ahead forecast would be the expectation that the value z will take on in t+1, given the information set available at t:, namely$E_t\{z_{t+1}|\Omega_t\} = B_1 z_t + ... + B_p z_{t+1-p}$, with $E\{e_{t+1}\} = 0$. Out-of-sample forecasts are easily obtained via forward iteration. In the case of a VAR(1): $z_t = B_1 z_{t-1} + e_t$, the optimal h periods ahead forecast would be

$$E_{t+h-1}\{z_{t+h}\} = B_1 z_{t+h-1} = B_1^h z_t$$

using the law of iterated expectations.

In practice, however, what is usually done is a "pseudo out-of-sample" or "rolling one-period-ahead" forecast in which the estimated VAR coefficients are updated each period to account for the most recently released data.[1] Assuming a data set of length T, one estimates the VAR model using data up to $T_1 < T$, then generates a one-period-ahead forecast and actual forecasting error for the current period, $T_1 + 1$. The estimation sample is then increased one observation

[1]See Marcellino, Stock, and Watson (2004).

to T_1+1 and a new one-period-ahead, this time for T_1+2, forecast is calculated using the new VAR estimates. The process is repeated until all the available data points are used to calculate a forecast error sequence from T_1+1 to T.

3.3.1 Forecast Evaluation

Forecasting performance is normally assessed using a user specified loss function, which is usually a non negative function of the forecast errors: $(\hat{z}_{t+h} - z_{t+h})$.[2] Most commonly, researchers select the model that either minimizes the mean squared forecast error $MSFE = \frac{1}{T-T_1}\sum_i^h(\hat{z}_{t+i} - z_{t+i})^2$ or the root mean squared error (RMSE) $\sqrt{\frac{1}{T-T_1}\sum_i^h(\hat{z}_{t+i} - z_{t+i})^2}$. The iterated multi-step-ahead forecasts obtained with a VAR are frequently compared with the ones generated from univariate AR(p) models and/or the random walk (constant) forecast for each individual variable of interest.

In practice, several statistics exist to determine whether the $MSFEs$ of two models are different, e.g., the F-statistic, $F = \sum_i^h(\hat{z}_{t+i} - z_{t+i})^2/\sum_i^h(\hat{z}_{t+i} - z_{t+i})^2$. However, the forecast errors must be normally distributed, serially uncorrelated, and contemporaneously uncorrelated with each for this expression to have an F distribution. Other statistics try to relax one or several of these assumptions, e.g., the Granger-Newbold and the Diebold-Mariano statistics.

3.3.2 Conditional Forecasts

Conditional forecasts are often calculated using VARs. To do so, one fixes a future trajectory of at least one of the variables in the VAR, thereby treating these variables as exogenous for forecasting purposes. For instance, researchers commonly assume paths for the interest rate, the price of oil and fiscal spending, among with other policy or exogenous variables, when producing a baseline forecast.

There are in fact several strategies available for implement conditional forecasts. For instance, Banbura, et al. (2015) propose the computation of the conditional forecast using the recursive period-by-period forecast technique of the Kalman filter while Waggoner and Zha (1999) use a Gibbs sampling algorithm to sample the VAR coefficients and thereby provide a distribution for the conditional forecasts.

3.3.3 Forecasting Using EViews

Unconditional VAR forecasts in EViews can be produced in two ways. The first approach relies on using the ***Forecast*** tab that is available once the reduced form VAR is estimated (Figure 3.1). The starting and ending date of the forecast is controlled by the sample period in the forecast dialog box, which is shown in the lower right-hand corner of Figure 3.1. EViews can calculate either static forecasts (i.e., forecasts based on actual data for the lagged variables) or dynamic

[2] A huge literature has developed around forecast evaluation. See West (2006).

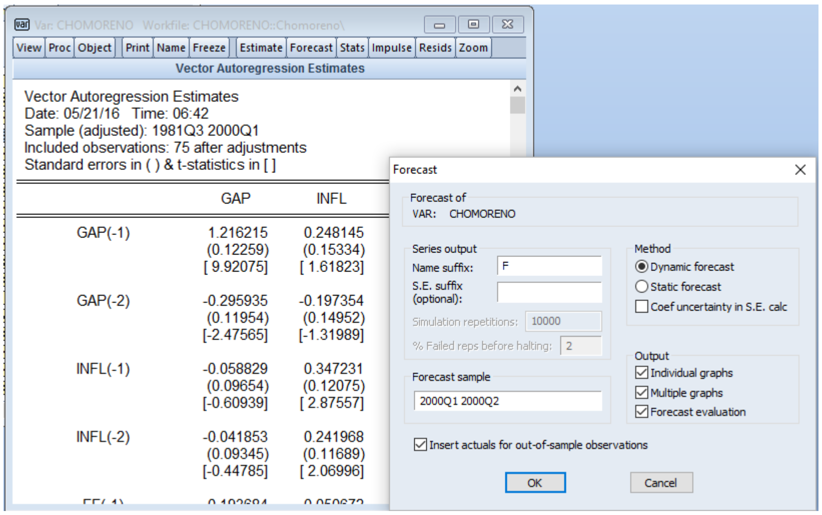


Figure 3.1: Forecasting a Reduced Form VAR using EViews: Direct Approach

forecasts. Dynamic forecasts use forecasted values for the lagged variables rather than actuals (assuming they are available) during the forecast horizon.

Figure 3.1 shows how to generate a two-period-ahead dynamic forecast beginning with the last date of the estimation period (2000Q1), thus permitting a forecast evaluation against actual data for 2000Q1. Pressing the **OK** button of this dialog box produces the output shown in Figure 3.2. Note that the forecasts are saved in the workfile using the same variables as in the VAR, but with (in this case) an F suffix (i.e., GAP_F, INFL_F, and FF_F) that can be controlled by the user.

The other approach to generating forecasts involves using the model simulator available in EViews. It is a far more flexible forecasting tool than the ***Forecast*** tab, allowing one to generate forecasts under alternative scenarios. In particular forecasts conditional upon an assumed path for one or more of the endogenous variables. The first step in using the model simulator is to create a model representation of the estimated VAR using ***the Proc→Make Model*** menu command (see Figure 3.3). Unconditional forecasts (e.g., for the small macro model) are produced by clicking on the ***Solve*** tab, which yields the baseline forecast of the model. The forecasts are stored ina workfile using the suffix "_0". The corresponding variable names are GAP_0, INFL_0, and FF_0. Forecasts with alternative scenarios, particularly for models with exogeneous variables, can be generated by defining a new scenario and specifying the time path of the control (typically exogeneous) variables under it.

It is also possible to exclude (and control) the time path of endogeneous variable in an alternative scenario, which yields a conditional forecast. Calculating a conditional forecast using the EViews model simulator must be done using an alternative scenario. The first step is to define the values of the con-

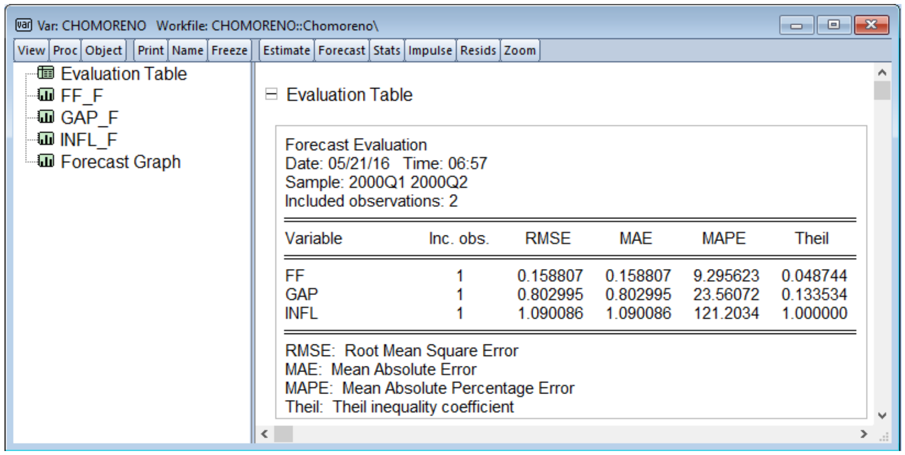


Figure 3.2: EViews Output: Forecast Tab

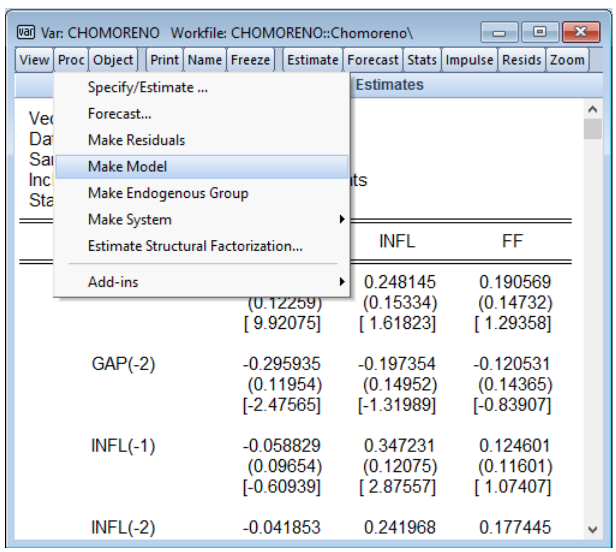


Figure 3.3: Creating a VAR model using EViews

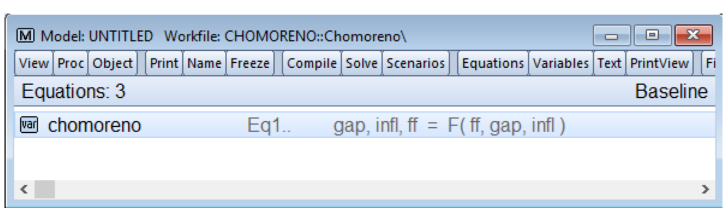


Figure 3.4: VAR Model Object: Chomoreno

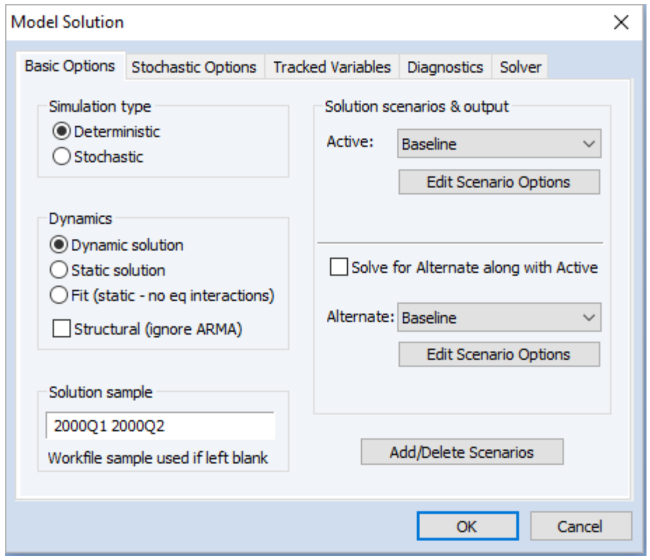
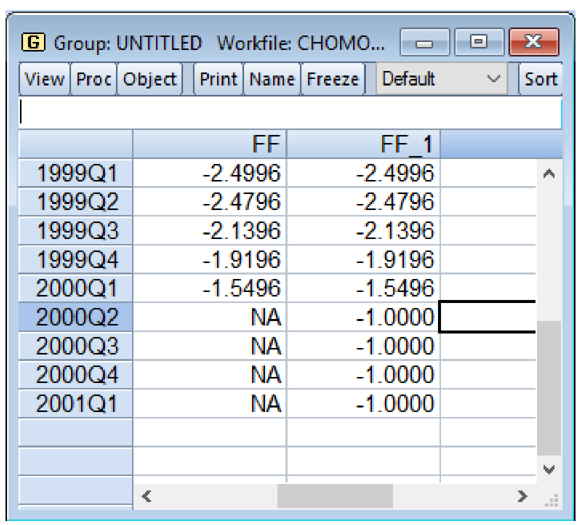


Figure 3.5: Generating a VAR Forecast Using the VAR Model Object



	FF	FF_1
1999Q1	-2.4996	-2.4996
1999Q2	-2.4796	-2.4796
1999Q3	-2.1396	-2.1396
1999Q4	-1.9196	-1.9196
2000Q1	-1.5496	-1.5496
2000Q2	NA	-1.0000
2000Q3	NA	-1.0000
2000Q4	NA	-1.0000
2001Q1	NA	-1.0000

Figure 3.6: Federal Funds Rate (FF) Under the Alternative Scenario (Scenario 1)

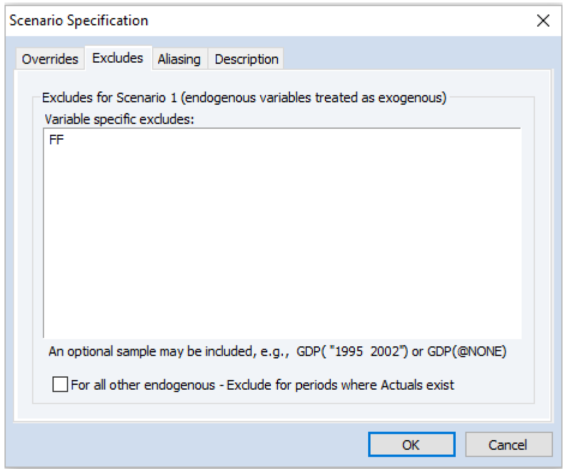


Figure 3.7: Conditional Forecasting Using the Eviews Model Simulator: Editing the Alternative Scenario

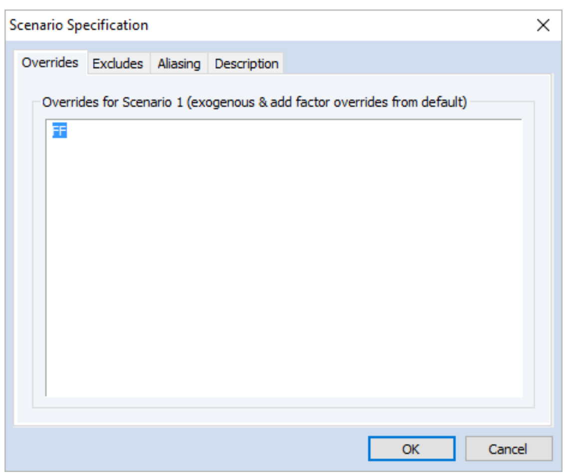


Figure 3.8: Conditional Forecasting Using the EViews Model Simulator: Overriding a Variable

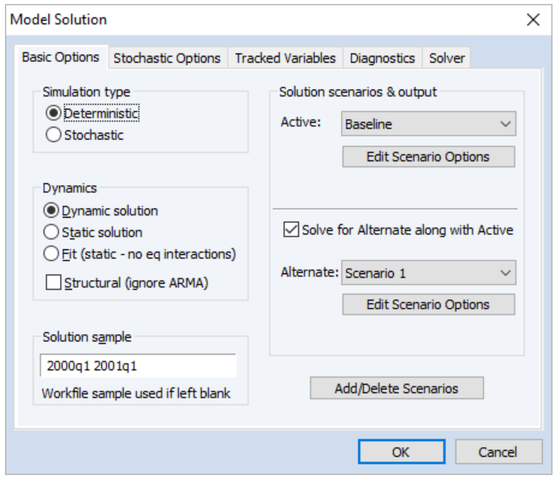
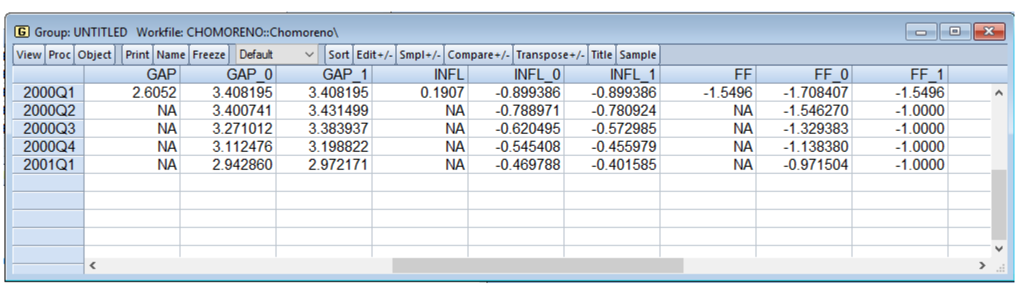


Figure 3.9: Simulating the Chomoreno model Under An Alternative Scenario



Group: UNTITLED Workfile: CHOMORENO::Chomoreno\

	GAP	GAP_0	GAP_1	INFL	INFL_0	INFL_1	FF	FF_0	FF_1
2000Q1	2.6052	3.408195	3.408195	0.1907	-0.899386	-0.899386	-1.5496	-1.708407	-1.5496
2000Q2	NA	3.400741	3.431499	NA	-0.788971	-0.780924	NA	-1.546270	-1.0000
2000Q3	NA	3.271012	3.383937	NA	-0.620495	-0.572985	NA	-1.329383	-1.0000
2000Q4	NA	3.112476	3.198822	NA	-0.545408	-0.455979	NA	-1.138380	-1.0000
2001Q1	NA	2.942860	2.972171	NA	-0.469788	-0.401585	NA	-0.971504	-1.0000

Figure 3.10: Conditional Forecasts for GAP and INFL using the Chomoreno VAR

ditioning variable during the forecast horizon. Assuming we are working with the Chomoreno model under "Scenario 1", and it is desired wish to condition upon the Federal Funds Rate (FF), the previous step amounts to setting the values of "FF_1" during the forecast horizon. For the purpose of illustration, we assume that the de-meaned Federal Funds Rate remains at -1.0000 during the forecast horizon 2000Q1-2001Q1 (see Figure 3.6). The next step is to exclude FF from the model simulation (thereby forcing FF to be an exogenous variable for simulation purposes), and then override its values during the forecast horizion. To do so in the EViews model simulator, click on the ***Solve*** tab and then ***Edit Scenario Options*** for ***Scenario 1***. This yields the dialog box shown in Figure 3.7. Click on the ***Exclude*** tab and insert "FF". To override the values of FF during the forecast horizon with those specified in FF_1, click on the **Overrides** tab, and then insert "FF" as shown in Figure 3.8. Lastly, solving the model under scenario 1 (Figure 3.7) produces conditional forecasts for the output gap and inflation that reflect the assumed values of FF during the forecast horizon. The resulting forecasts under the baseline (unconditional forecast) and alternative (conditional) forecast for GAP and INFL are shown in Figure 3.10.[3]

3.4 Bayesian VARs

VARs often involve estimating a large number of coefficients compared to the available number of observations, resulting in imprecisely estimated coefficients (the "over-fitting" problem). Whilst this may not be too important for the estimation of impulse responses it can result in extremely bad forecasts. Parsimonious models tend to be better at forecasting. For this reason one might wish to restrict the number of parameters being estimated in some way. One way is to omit lagged values of variables in some equations, i.e. not to keep order p lags in every equation of the VAR. A literature has emerged on good ways of determining the lag structure in individual equations that is often referred to as "best sub-set VARs". This is not available in EViews. Instead the simplification used in EViews involves applying Bayesian methods that impose useful prior distributions upon the complete set of VAR coefficients so as to acheive parsimony. Hence, it is possible to adopt a Bayesian VAR (BVAR) by utilizing the VAR menu, as the screen shot in Figure 3.11 shows.

The reason that BVARs may be effective in forecasting is that the priors relating to the VAR coefficients involve far fewer parameters than the original set in the $VAR(p)$ and they also impose some quantitative constraints that rule out certain parts of the parameter spece. The priors need not be correct. As has been often demonstrated, bad models (in terms of their economic rationale) can win forecasting competitions. However, to be successful the priors should impose some structure upon the VAR which reflects the nature of the data.

[3] See the program files *forecast.prg* and *forecast_rolling.prg* for complete examples of how to implement the above using the EViews command line language.

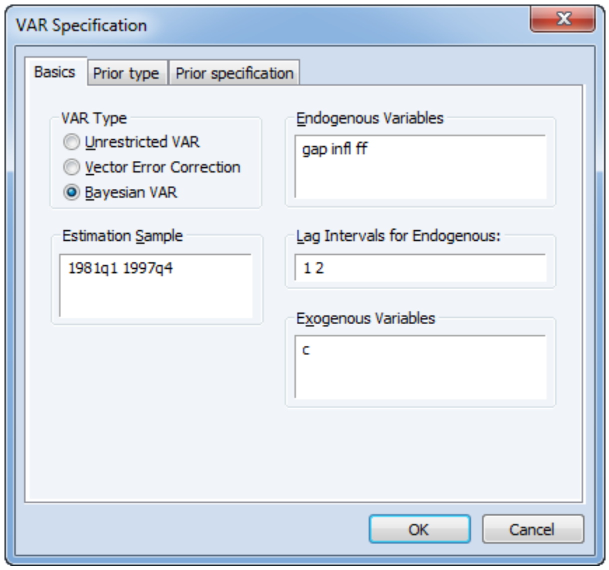


Figure 3.11: Estimating a Bayesian VAR using the EViews VAR Object

Often this is done very loosely. Thus, the very first method used for producing BVARs was that of Litterman (1986). It had been noted for many years that time series in macroeconomics and finance tended to be very persistent. Hence, looking at a VAR with $n = 2$ and $p = 1$, the first equation in the system would be $z_{1t} = b^1_{11} z_{1t-1} + b^1_{12} z_{2t-1} + e_{1t}$ so that persistence would mean b^1_{11} would be close to unity. In contrast b_{12} was likely to be zero. Hence Litterman used this in formulating a prior on the VAR coefficents. This prior is now incorporated into EViews and can be invoked by clickiing on the **prior type** tab shown in Figure 3.11. Because some other priors are also available it is necessary to begin with some general discussion relating to Bayesian methods and BVARs.

Consider a standard regression model with unknown coefficients β and an error variance-covariance matrix Σ_e, i.e.

$$z_t = x_t'\beta + e_t,$$

where x_t includes the lags of z_t, any exogeneous variables in the system, and

$$e_t \sim nid\ N(0, \Sigma_e).$$

Given a prior distribution for β conditional on Σ_e - $p(\beta|\Sigma_e)$ - Bayes' theorem is used to combine the likelihood function of the data with the prior distribution of the parameters to yield the posterior distribution of β, viz:

$$\underbrace{p(\beta|\Sigma_e, Z)}_{posterior} = \frac{\overbrace{L(Z|\beta, \Sigma_e)}^{Likelihood}\overbrace{p(\beta|\Sigma_e)}^{Prior}}{p(Z)}, \quad (Bayes\ theorem)$$

where $p(Z) = \int p(z|\beta)p(\beta)d\beta$ is a normalizing constant. It follows that the posterior distribution is proportional to the likelihood function times the prior distribution:

$$\underbrace{p(\beta|\Sigma_e, Z)}_{posterior} \propto \overbrace{L(Z|\beta, \Sigma_e)}^{Likelihood}\overbrace{p(\beta|\Sigma_e)}^{Prior}.$$

Now to prepare forecasts we would need an estimate of β. One estimate would be the mode of the posterior forβ , and this can be found by maximizing

$$C(\beta) = ln\{L(Z|\beta, \Sigma_e)\} + ln(p(\beta|\Sigma_e). \quad (3.1)$$

There are other possible estimates forβ, e.g. the mean of the posterior. In the event that the posterior is normal then the mode and mean will correspond but they can be different in other cases. If one is happy with using the mode than it is only necessary to maximize $C(\beta)$ rather than finding a complete posterior density.

Although Bayesian methods today can find posteriors for a number of different types of priors by simulation methods, when BVARs were first proposed it was more common to select priors in such a way as to obtain a closed form solution for the posterior distribution. This led to what were termed natural conjugate priors, i.e., priors which in combination with the likelihood would produce a tractable posterior, generally having the same density as the prior. In most instances the prior was made normal and the posterior was as well. For instance, if the prior for β in the regression model above is assumed to be normally distributed

$$p(\beta) \sim N(\underline{b},\ \underline{V}),$$

then the posterior will also be normal. In particular, the mode and mean estimate of β would be a matrix weighted average of the OLS estimates and the researcher's priors:

$$\overline{b} = [(\underline{V}^{-1} + \Sigma_e^{-1} \otimes (X'X)]^{-1}[(\underline{V}^{-1}\underline{b} + (\Sigma_e^{-1} \otimes X')y].$$

It is clear from this formula that BVARs will tend to shrink the estimated coefficients of the VAR model towards the prior mean and away from the OLS estimates, and it is this which can give the prediction gains.

A variety of Bayesian priors have been developed specifically for VARs, and we now review two of the most popular priors used in applied research. Note from the formula for $\overline{b}$ above that it depends on Σ_e. One needs to produce some estimate of Σ_e and this might be done either using OLS type information or by producing a Bayesian estimate of Σ_e. In the latter case we would need a prior for it. We will review the two priors used in EViews that relate specifically to

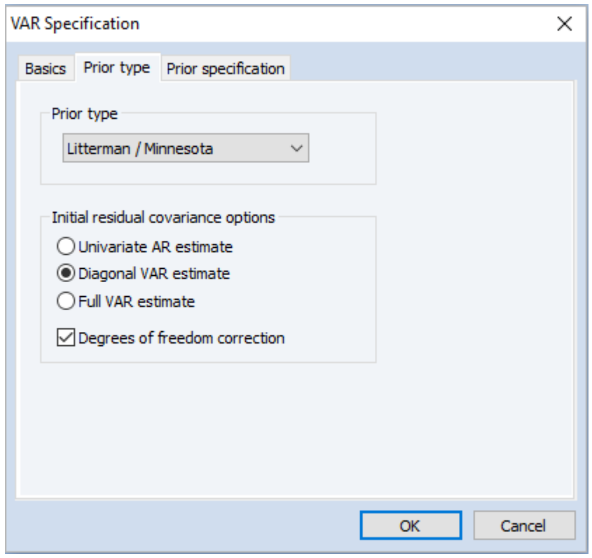


Figure 3.12: Bayesian VAR Estimation in EViews: "Prior type" Dialog Box

VARs - the Minnesota prior and the Normal-Wishart prior. In chapter 4 we will look at other alternatives based on Sims and Zha (1998) that focus on SVARs.

3.4.1 The Minnesota prior

It is worth starting with the EViews screen that shows the initial set-up for BVARs. This is given in Figure 3.12 and, as we have said previously, one of the choices of prior is that referred to as Litterman/ Minnesota. This prior for β is normal conditional upon Σ_e. Hence some assumption is needed about the nature of Σ_e and how it is to be estimated, and this accounts for the three choices in the box of Figure 3.12. These involve selecting one of the following: (a) Use estimates of the residual variances from fitting AR(1) models to each series (b); Assume that Σ_e is replaced by an estimate, $\hat{\Sigma}_e$, in which the diagonal elements σ_i^2 correspond to the OLS estimated VAR error variances; and (c) Estimate the complete Σ_e implied by the VAR (the *df* argument controls whether the initial residual covariance is to be corrected for the available degrees of freedom). One reason for not using (c) in early studies was that the estimated matrix might be singular, since there may not have been enough observations when n and p were large.

Once a decision has been made about how Σ_e is to be handled one can then proceed to describe to EViews what the prior for β is. This is done in the screen shot in Figure 3.13 by reference to a set of hyper-parameters. These are the standard options, although it would also be possible for the user to specify $\underline{b}$

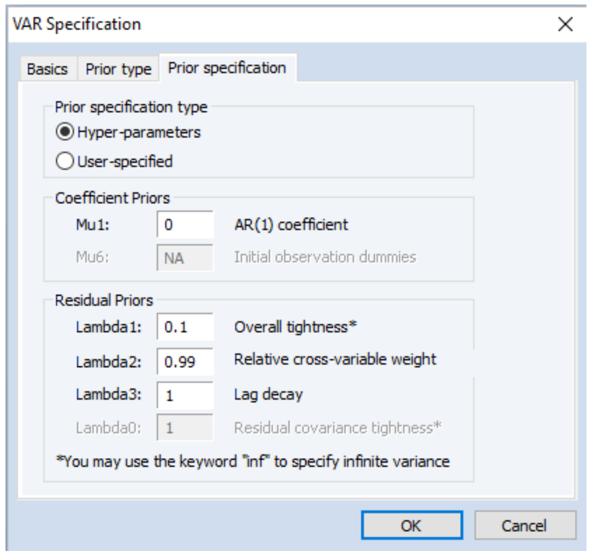


Figure 3.13: Bayesian VAR Estimation in EViews: "Prior specification" Dialog Box

and $\underline{V}^{-1}$directly. For the automatic options the vector of prior means $\underline{b}$ are all the same, being given by the value of the parameter μ_1. In most instances we would want $\mu_1 = 1$ or something close to unity in order to capture the persistence in economic time series. However, this would not be true if the data z_t was say GDP growth. Then we would want the prior mean of lagged growth to be either zero or a small number. In Figure 3.13 the parameter *Lambda1* (i.e., λ_1) controls the overall tightness of the prior for β, and should be close to zero if there is more certainty about the prior. We set *Lambda1* to 0.1, implying a relatively strong prior.[4] *Lambda2* (λ_2) controls the importance of the lagged variables of the j'th variable in the i'th equation ($i \neq j$) of the VAR (these are termed cross-variable weights in Figure 3.13). λ_2 must lie between 0 and 1. When *Lambda2* is small, the cross-lag variables in the model play a smaller role in each equation. Lastly, *Lambda3* (i.e., λ_3) determines the lag decay rate via $l^{\lambda 3}$ where l is the lag index. We set this parameter to 1 (unity) for no decay. Note that since this hyper-parameter also appears in the denominator of the expression for the prior variances of the coefficients of the cross-lag variables $\left(\frac{\lambda_1 \lambda_2 \sigma_i}{l^{\lambda 3} \sigma_j}\right)^2$, then the diagonal element for the second lag will be $\left(\frac{\lambda_1 \lambda_2 \sigma_i}{2 \sigma_j}\right)^2$, and so on for higher order lags (if any).

The Minnesota prior is specifically designed to center the distribution of β so that each variable behaves as a random walk (see Del Negro and Schorfheide

[4]Set *Lambda1* to 10 or higher for a non-informative (more uncertain) prior. In this case, the estimated parameters will be close to the unrestricted VAR coefficients.

(2010)). The prior was chosen because random walks are usually thought to be good predictors of macroeconomic time series.

For illustrative purposes, consider the following bi-variate VAR:

$$\begin{aligned}\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} &= \begin{pmatrix} \beta_{11}^1 & \beta_{12}^1 \\ \beta_{21}^1 & \beta_{22}^1 \end{pmatrix}\begin{pmatrix} z_{1t-1} \\ z_{2t-1} \end{pmatrix} \\ &+ \begin{pmatrix} \beta_{11}^2 & \beta_{12}^2 \\ \beta_{21}^2 & \beta_{22}^2 \end{pmatrix}\begin{pmatrix} z_{1t-2} \\ z_{2t-2} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}. \end{aligned} \tag{3.2}$$

We also assume that the covariance matrix of the population errors is diagonal, i.e. option 2 in Figure 3.12.

$$\Sigma_e = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

The Minnesota prior assumes that the prior means of β_{11}^1 and β_{22}^1 are unity and all other coefficients have a mean of zero. The prior for the variance-covariance matrix of the coefficients can be represented as:

$$\left(\frac{\lambda_1}{l^{\lambda_3}}\right)^2 \; for \quad (i = j) \tag{3.3}$$

$$\left(\frac{\lambda_1 \lambda_2 \sigma_i}{l^{\lambda_3} \sigma_j}\right)^2 \; for \quad (i \neq j), \tag{3.4}$$

where σ_i^2 is the i-th diagonal element of Σ_e.

It is evident that the hyper parameters in the covariance matrix, $(\lambda_1, \lambda_2, \lambda_3)$ influence the estimated coefficients as follows:

1. λ_1 controls the prior standard deviation of β_{11}^1 and β_{22}^1. These parameters correspond to the first lag of the first variable z_{1t} in the first equation and the first lag of the second variable z_{2t} in the second equation. In the general case of n variables, the smaller λ_1 is the more the first-lag coefficients β_i^1 $i = 1, ..., n$ will shrink toward unity, while the remaining lag coefficients will shrink more towards zero.

2. λ_2 controls the variance of the coefficients of variables that are different from the dependent variable of the i^{th} equation. Those coefficients move closer to zero as λ_2 declines.

3. λ_3 influences the estimated coefficients on lags beyond the first. As λ_3 increases, the coefficients on higher order lags shrink toward zero.

Note that the ratio (σ_i/σ_j) in (3.4) is included to account for differences in the units of measurement of the variables.

The posterior of the Minnesota prior has a closed form solution and Koop and Korobilis (2010) highlight that a key advantage of it is that the posterior is in fact a Normal distribution. Several variants of this Minnesota prior have been used in applied research, including one that uses a non-diagonal error covariance matrix, different ways of introducing the lag decay, and a different way of introducing the priors through dummy variables (see Theil and Goldberger (1961) and Del Negro and Schorfheide (2010) for specific examples). We will look at the latter later in this section.

3.4.1.1 Implementing the Minnesota prior in EViews

We now demonstrate how to estimate a VAR (BVAR) in EViews using the Minnesota prior. The small macro model of chapter 2 is estimated using the available data to 1997Q1, following which out-of-sample forecasts are generated for 1998Q1 - 2000Q1. The BVAR can be estimated using the standard EViews VAR object (see Figure 3.11). Given the persistence in the series, we will assume that *gap* and *ff* are I(1) while *infl* is I(0). After choosing the BVAR option it is necessary to specify the sample period, the desired number of lags, and the variables in the system (in this case, the differences in *gap* and *ff* - *dgap* and *dff*- as well as the level of inflation *infl).* Following these data transformations both the prior and the values of the associated hyper-parameters need to be selected by clicking on the "Prior Type" (Figure 3.12) and then the "Prior Specification" (Figure 3.13) tabs respectively.[5] Coefficient estimates for the VAR and BVAR are shown in Figure 3.14. Compared to the standard OLS estimates, the most discernable difference is that the BVAR estimates for the first own lag of each variable are significantly smaller (as expected, given the setting of *Mu1* = 0). We will see the impact this has on the model's forecasting accuracy below.

While the Minnesota prior is still actively used because of its success as a forecasting tool, it ignores any uncertainty associated with the variance-covariance matrix Σ_e. This assumption is relaxed by use of a Normal-Wishart prior system, which we now review and demonstrate.

3.4.2 Normal-Wishart prior

Rather than just replacing Σ_e by some estimate from the data it might be desired to estimate it with Bayesian methods. A natural conjugate prior for the covariance matrix Σ_e- $p(\underline{\Sigma}_e^{-1})$ - is the Wishart distribution and, with the prior for β being normal, this yields a posterior for β that is a product of a Normal and a Wishart distribution. The prior for β depends on $\underline{\beta}$, $\underline{V}$ while that for Σ_e depends on two parameters ν, $\underline{S}$. $\underline{V}$ depends on a hyper-parameter $\lambda_1$just as it did with the Minnesota prior. EViews sets ν to equal the degrees of freedom and $\underline{S}$ to the identity matrix, so that only two parameters need to be set by the user. Basically the two parameters are set in much the same way as for the Minnesota prior. Because there are now less hyper-parameters involved in the prior β some restrictions will be implied, specifically that the prior covariance of

[5] The corresponding command to estimate the BVAR in an EViews program is:
var chobvar.bvar(prior=lit, initcov=diag,df,mu1=0,L1=0.1,L2=0.99,L3=1) 1 2 d(gap) infl d(ff) @ c

Var: CHOMORENO Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Vector Autoregression Estimates

Vector Autoregression Estimates
Date: 06/15/16 Time: 09:28
Sample (adjusted): 1981Q4 2000Q1
Included observations: 74 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	D(GAP)	INFL	D(FF)
D(GAP(-1))	0.318776	0.316152	-0.095529
	(0.11554)	(0.16249)	(0.14534)
	[2.75905]	[1.94568]	[-0.65730]
D(GAP(-2))	0.248097	-0.260978	0.638923
	(0.10994)	(0.15462)	(0.13830)
	[2.25664]	[-1.68789]	[4.61996]
INFL(-1)	-0.070958	0.472602	-0.093315
	(0.08076)	(0.11357)	(0.10158)
	[-0.87867]	[4.16122]	[-0.91860]
INFL(-2)	-0.061297	0.306842	-0.015323
	(0.07826)	(0.11006)	(0.09844)
	[-0.78329]	[2.78802]	[-0.15566]
D(FF(-1))	0.105286	-0.053422	0.178995
	(0.08636)	(0.12146)	(0.10864)
	[1.21913]	[-0.43985]	[1.64766]
D(FF(-2))	-0.312247	0.221139	-0.355217
	(0.08227)	(0.11571)	(0.10349)
	[-3.79526]	[1.91122]	[-3.43231]
C	-0.037879	-0.097685	-0.195668
	(0.06583)	(0.09259)	(0.08281)
	[-0.57536]	[-1.05505]	[-2.36274]

Var: CHOMORENO Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Bayesian VAR Estimates

Bayesian VAR Estimates
Date: 06/15/16 Time: 11:14
Sample (adjusted): 1981Q4 2000Q1
Included observations: 74 after adjustments
Prior type: Litterman/Minnesota
Initial residual covariance: Diagonal VAR
Hyper-parameters: Mu: 0, L1: 0.1, L2: 0.99, L3: 1
Standard errors in ( ) & t-statistics in [ ]

	D(GAP)	INFL	D(FF)
D(GAP(-1))	0.180298	0.053327	0.041792
	(0.07177)	(0.10041)	(0.08982)
	[2.51221]	[0.53110]	[0.46529]
D(GAP(-2))	0.052499	-0.060763	0.108425
	(0.04459)	(0.06220)	(0.05563)
	[1.17749]	[-0.97686]	[1.94887]
INFL(-1)	-0.095459	0.413226	-0.085798
	(0.04387)	(0.06199)	(0.05518)
	[-2.17614]	[6.66568]	[-1.55496]
INFL(-2)	-0.033091	0.130396	-0.034612
	(0.03042)	(0.04312)	(0.03826)
	[-1.08796]	[3.02405]	[-0.90469]
D(FF(-1))	0.066021	-0.034680	0.073372
	(0.05694)	(0.08007)	(0.07199)
	[1.15943]	[-0.43310]	[1.01915]
D(FF(-2))	-0.044235	0.032808	-0.047940
	(0.03519)	(0.04949)	(0.04462)
	[-1.25710]	[0.66298]	[-1.07444]
C	-0.001504	-0.137810	-0.172853
	(0.06349)	(0.08929)	(0.07988)
	[-0.02369]	[-1.54336]	[-2.16393]

Figure 3.14: VAR versus BVAR Estimates (Minnesota Prior): Chomorono Model, 1981Q3-1997Q4.

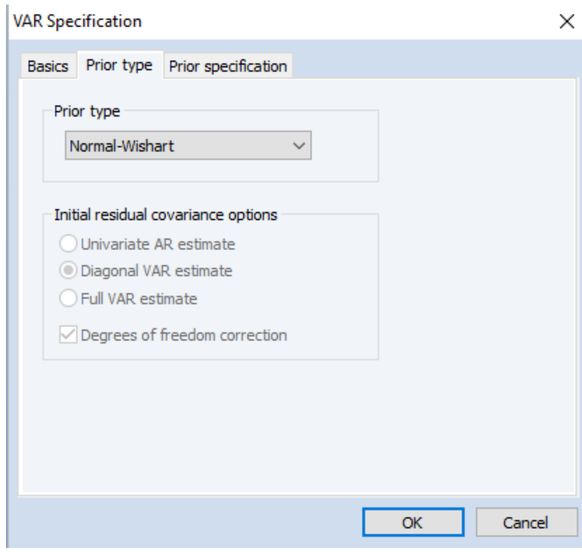


Figure 3.15: Selecting the Normal-Wishart Prior in EViews

coefficients of different equations are proportional.[6] For instance, in the VAR(2) example above, any reduction in the prior variance of β_{11}^1 results in a reduction of the variance of β_{21}^1 as well.

3.4.2.1 Implementing the Normal-Wishart prior in EViews

As in the previous example, Bayesian estimation with a Normal–Wishart prior can be carried out in EViews using the standard VAR object and setting the prior type to "Normal-Wishart" (see Figure 3.15). [7].

The next step is to set the values of the hyper-parameters, which are similiar to those for the Minnesota prior (see Figure 3.16), but which influence the parameter estimates differently. The *Mu1* parameter governs the prior concerning the properties of the time-series process. It should be set to zero (or very small) when the modeler believes that the series in the VAR are stationary, and unity if the series in the VAR are thought to be better modeled with a unit root process. Because we have differenced the two series that seem close to having unit root characteristics - *gap* and *ff* - we set $\mu_1 = 0$. The remaining hyper-parameter that EViews supports for the Normal-Wishart prior is *Lambda1*, the overall tightness parameter. A large value for *Lambda1* implies greater certainty about the prior for β, which is exactly the opposite of how *the Lambda1* parameter works in the

[6]See Gonzalez 2016.
[7]The corresponding command line code is
var chobvar1.bvar(prior=nw, df, mu1=0.01, L1=10) 1 2 d(gap) infl d(ff) @ c

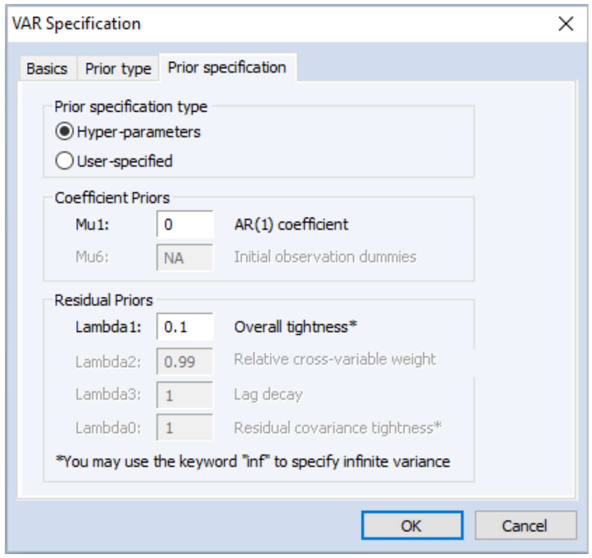


Figure 3.16: Specifying the Hyper-Parameters for the Normal-Wishart Prior

case of the Minnesota prior.

The parameter estimates for β corresponding to these settings of the hyper parameter are presented in Figure 3.17. Again, relative to the OLS parameter estimates, the own lag coefficients have shrunk, but not as much as in the case of the Minnesota prior.

3.4.3 Additional priors using dummy observations or pseudo data

The use of dummy observations as a way of introducing *a priori* knowledge about regression coefficients dates back to Theil and Goldberger (1960). They write the *a priori* information as pseudo data and then estimate an augmented regression including this data. Although it does not seem to have had great use with VARs after the initial work by Litterman, it does appear in connection with the priors of Sims and Zha (1998) that will be discussed in Chapter 4, and so it is convenient to discuss the method in the VAR context.

Suppose that the researcher believes that the single coefficient β in the regression $z = X\beta + e$ should be 0.6, with a standard deviation of 0.2. This prior information can be introduced via the pseudo data $r = R\beta + v$, where uncertainty about β is captured by $v \sim N(0, \Phi)$. In terms of the information given earlier the pseudo-data would be written as $\overbrace{(0.6)}^{r} = \overbrace{(1)}^{R}\,\overbrace{(\beta)}^{\beta} + v$ and $\Phi = .04$, and added on as an additional observation in the data set. Thus the augmented

Var: CHOMORENO Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Bayesian VAR Estimates

Bayesian VAR Estimates
Date: 06/15/16 Time: 09:36
Sample (adjusted): 1981Q4 2000Q1
Included observations: 74 after adjustments
Prior type: Normal-Wishart
Hyper-parameters: Mu: 0, L1: 0.1
Standard errors in ( ) & t-statistics in [ ]

	D(GAP)	INFL	D(FF)
D(GAP(-1))	0.317622	0.314104	-0.094176
	(0.11611)	(0.15941)	(0.14352)
	[2.73560]	[1.97037]	[-0.65619]
D(GAP(-2))	0.247294	-0.259331	0.635783
	(0.08122)	(0.11152)	(0.10040)
	[3.04471]	[-2.32550]	[6.33264]
INFL(-1)	-0.071020	0.471837	-0.092840
	(0.08689)	(0.11929)	(0.10740)
	[-0.81740]	[3.95527]	[-0.86444]
INFL(-2)	-0.061460	0.307209	-0.016020
	(0.11050)	(0.15172)	(0.13660)
	[-0.55618]	[2.02482]	[-0.11728]
D(FF(-1))	0.105270	-0.052917	0.178325
	(0.07870)	(0.10806)	(0.09729)
	[1.33755]	[-0.48970]	[1.83300]
D(FF(-2))	-0.311149	0.220656	-0.353903
	(0.08279)	(0.11367)	(0.10234)
	[-3.75824]	[1.94117]	[-3.45815]
C	-0.037652	-0.097567	-0.195347
	(0.06628)	(0.09100)	(0.08193)
	[-0.56808]	[-1.07215]	[-2.38436]

Figure 3.17: Bayesian VAR Estimates using a Normal-Wishart Prior

"observations" would be

$$\begin{pmatrix} z \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} e \\ v \end{pmatrix}.$$

Estimation by OLS then yields the posterior for $\beta$of $\beta|z \sim N(\overline{\beta}, \overline{V_\beta})$, where

$$\overline{\beta} = ((X'\Sigma_e^{-1}X) + (R'\Phi^{-1}R))^{-1}(X'\Sigma_e^{-1}z + (R'\Phi^{-1}r))$$

$$\overline{V} = s^2((X'\Sigma^{-1}X) + (R'\Phi^{-1}R))^{-1}$$

Now this idea of augmenting the data set with pseudo-data capturing the prior information can be used in many ways. Two important uses of it have been to account for what are described in EViews as "sum of coefficients dummies" and "initial observation dummies"

3.4.3.1 Sum of Coefficients Dummy Prior

Suppose we had a VAR(2) in two variables. Then the first equation would be

$$z_{1t} = b_{11}^1 z_{1t-1} + b_{11}^2 z_{1t-2} + b_{12}^1 z_{2t-1} + b_{12}^2 z_{2t-2} + e_{1t}. \tag{3.5}$$

Now it might not make sense to impose the Minnesota prior that puts the prior mean of b_{11}^1 to unity. Insead we might want to impose b_{11}^1+ b_{11}^2 as having a prior of unity. To capture this we would define the pseudo-data as

$$\mu_5 s_1 = \mu_5 s_1 b_{11}^1 + \mu_5 s_1 b_{11}^2 + v_1,$$

where s_1 is some quantity that reflects the units of y_1 e.g. a mean or standard deviation of y_1 over some sub-sample. Hence $r = \mu_5 s_1$and $R = [\ \mu_5 s_1 \quad \mu_5 s_1 \quad 0 \quad 0\ ]$. Using the pseudo-data then implies that

$$1 = b_{11}^1 + b_{11}^2 + (\mu_5 s_1)^{-1} v_1,$$

and we see that the sum of coefficents restriction will hold as $\mu_5 \to$infinity.

3.4.3.2 The Initial Observations Dummy Prior

Consider an SVAR(1) with 2 variables. It has the form

$$z_t = B_1 z_{t-1} + e_t, \tag{3.6}$$

and can be written as

$$z_{1t} = b_{11}^1 z_{1t-1} + b_{12}^1 z_{2t-1} + e_{1t} \tag{3.7}$$

$$z_{2t} = b_{21}^1 z_{1t-1} + b_{22}^1 z_{2t-1} + e_{2t}. \tag{3.8}$$

Then

$$\Delta z_{1t} = (b^1_{11} - 1)z_{1t-1} + b^1_{12}z_{2t-1} + e_{1t} \tag{3.9}$$

$$\Delta z_{2t} = b^1_{21}z_{1t-1} + (b^1_{22} - 1)z_{2t-1} + e_{2t}, \tag{3.10}$$

which has the form

$$\Delta z_t = Dz_{t-1} + e_t. \tag{3.11}$$

Now suppose we define the two pieces of pseudo-data for the SVAR(1) as

$$[\mu_6\overline{y}_1\mu_6\overline{y}_2] = [\mu_6\overline{y}_1\mu_6\overline{y}_2]\begin{bmatrix} b^1_{11} & b^1_{12} \\ b^1_{21} & b^1_{22} \end{bmatrix} + \begin{bmatrix} v_1 & v_1 \end{bmatrix}. \tag{3.12}$$

It will imply that

$$(1 - b^1_{11}) = \frac{\overline{y}_2}{\overline{y}_1}b^1_{21} + \frac{\nu_1}{\mu_6\overline{y}_1}$$

and

$$(1 - b^1_{22}) = \frac{\overline{y}_1}{\overline{y}_2}b^1_{12} + \frac{\nu_2}{\mu_6\overline{y}_2}.$$

Now, as $\mu_6 \to \infty$, we see that these constraints become

$$(1 - b^1_{11}) = \frac{\overline{y}_2}{\overline{y}_1}b^1_{21}$$

and

$$(1 - b^1_{22}) = \frac{\overline{y}_1}{\overline{y}_2}b^1_{12}. \tag{3.13}$$

Eliminating the ratios of $\overline{y}_1$ and $\overline{y}_2$ we find that $(1 - b^1_{11})((1 - b^1_{22}) - b^1_{12}b^1_{21} = 0$, i.e. the matrix D is singular. Hence it can be written as $\gamma\delta'$, where δ' can be expressed in terms of any co-integrating vector between z_{1t} and z_{2t}. Thus the use of the pseudo-data, along with allowing $\mu_6 \to \infty$, implies co-integration between the two variables.

Instead of implementing a co-integration constraint EViews imposes a co-trending constraint. Here the model has the form

$$\Delta z_{1t} = (b^1_{11} - 1)z_{1t-1} + b^1_{12}z_{2t-1} + c_1 + e_{1t} \tag{3.14}$$

$$\Delta z_{2t} = b^1_{21}z_{1t-1} + (b^1_{22} - 1)z_{2t-1} + c_2 + e_{2t}. \tag{3.15}$$

Imposing a prior on D of zero means that $c_1$and c_2 will be the deterministic trends in each series. For there to be a common one $c_1 = c_2$. To impose this restriction with dummy variables we will need R to have the row $\begin{bmatrix} \mu_6\overline{y}_1 & \mu_6\overline{y}_2 & \mu_6 \end{bmatrix}$ and r will be $\begin{bmatrix} \mu_6\overline{y}_1 & \mu_6\overline{y}_2 \end{bmatrix}$, as β will now involve the constant term $c_1 = c_2 = c$. In this instance $\mu_6 \to \infty$ implies co-trending between the two variables but not co-integration. To impose the latter would require an extra constraint reflecting the fact that the co-integrating and co-trending vectors need not be the same.

Table 3.1: Forecasting Performance of the Small Macro Model using Bayesian Estimation Methods, 1998:1-2000:1

Prior	Variable	RMSE	MAE
Standard VAR	$infl$	0.8427	0.8208
	Gap	0.6029	0.4844
Minnesota	$infl$	1.2719	1.2209
	Gap	1.3475	1.0856
Normal-Wishart	$infl$	0.8439	0.8218
	Gap	0.6033	0.4848

3.4.4 Forecasting with Bayesian VARs

Lastly, notwithstanding the fact that the small macro model has a relatively small number of estimated parameters, we used the the VAR and the two BVAR models estimated above to perform an out-of-sample forcasting experiment for the 1998Q1-2000q1 period (9 quarters), focusing on inflation ($infl$) and the output gap (gap). The results are given in Table 3.1. The results suggest little if any gain from using Bayesian methods. Also, the poor performance of the Minnesota prior relative to the unrestricted VAR suggests that the we have imposed a poor prior on the model.

3.5 Computing Impulse Responses

It is rarely the case that one is interested in the B_j. For this reason Sims (1980) suggested that one change the focus to how the shock e_{kt} would impact upon z_{jt}, i.e. to ask what is the response of $z_{j,t+M}$ to a shock e_{kt}? Accordingly, it is the partial derivative $\frac{\partial z_{j,t+M}}{\partial e_{kt}}$ that is of interest. These partial derivatives were called *impulse responses* since they showed the response of the variable z_j M periods ahead from t to a temporary one unit change in e_{kt}, i.e. the latter was raised by one unit at t but then set back to its normal value for $t+1, ..., t+M$. If variables (x_t) are exogenous then $\frac{\partial z_{j,t+M}}{\partial x_{kt}}$ were called *dynamic multipliers,* so impulse responses $\frac{\partial z_{j,t+M}}{\partial e_{kt}}$ are the equivalent concept once one views the e_{kt} as the exogenous variables. EViews produces these but not dynamic multipliers.

To examine the computation of response functions more carefully, and to flag many issues considered later, take a single variable following an AR(1) process:

$$\begin{aligned}
z_{1t} &= b_{11} z_{1t-1} + e_{1t} \\
&= b_{11}(b_{11} z_{1t-2} + e_{1t-1}) + e_{1t} \\
&= e_{1t} + b_{11} e_{1t-1} + b_{11}^2 z_{1t-2} \\
&= e_{1t} + b_{11} e_{1t-1} + b_{11}^2 e_{1t-2} + b_{11}^3 e_{1t-3} + ... \\
\therefore z_{1t+M} &= e_{1t+M} + ... + b_{11}^M e_{1t} + ...
\end{aligned}$$

and, provided $|b_{11}| < 1$ (*stationarity* holds), the term involving the initial e_{1t} will disappear as $M \to \infty$, as it has weight b_{11}^M. From this it is clear that the impulse responses are 1 $(M = 0)$, $b_{11}(M = 1), ...b_{11}^j(M = j)$...

To get a generalization of the simple treatment above to systems, the impulse responses D_l can be regarded as the weights attached to e_t in a *Moving Average (MA) representation for* z_t. Hence, when z_t is an $n \times 1$ vector, the MA will be

$$z_t = D_0 e_t + D_1 e_{t-1} + D_2 e_{t-2} + ...$$

Using the lag operator $L^k z_t = z_{t-k}$ the VAR can be written as

$$B(L)z_t = (I_n - B_1 L - ... - B_p L^p)z_t = e_t,$$

where I_n is the identity matrix of dimension n. Therefore $z_t = B^{-1}(L)e_t$, making $D(L) = B^{-1}(L)$ and $B(L)D(L) = I_n$. So, if z_t follows a VAR(1), $B(L) = I_n - B_1 L$ and therefore

$$\begin{aligned}(I - B_1 L)(D_0 + D_1 L + D_2 L^2 + ...) &= D_0 + (D_1 - B_1 D_0)L + (D_2 - B_1 D_1)L^2 + .. \\ &= I_n.\end{aligned}$$

Grouping and equating powers of L on the LHS and RHS gives

$$\begin{aligned}D_0 &= I \\ D_1 &= B_1 D_0 = B_1 \\ D_2 &= B_1 D_1 = B_1^2 \\ &. \\ &.\end{aligned}$$

Note that, since B_1 is a matrix, the $(i, j)'th$ element of D_M will be the impulse responses of z_{jt+M} to e_{kt}, i.e. the *j'th* variable to the $k'th$ shock, rather than just z_{jt+M} to e_{jt} as was the situation in the one variable case.

In general, for a VAR(p) the *impulse response function*, designated $D_l = \frac{\partial z_{t+l}}{\partial e_t}$, can be found by recursively solving

$$D_l = B_1 D_{l-1} + ... + B_p D_{l-p},$$

with the initial conditions for $l = 0, ..., p$ having to be determined. In all cases $D_0 = I_n$ provides initial values. When $p = 1$, application of the recursive equation gives $D_0 = I_n; D_1 = B_1; D_2 = B_1 D_1 = B_1^2$... When $p = 2$ the recursions yield $D_0 = I, D_1 = B_1 D_0, D_2 = B_1 D_1 + B_2 D_0$ etc.

To get impulse responses for the small macro model estimate the VAR(2) and then click on the commands ***View*** $\to$ ***Impulse Response***. The screen in Figure 3.18 will appear. It is clear from it that some decisions need to be made. First, it is necessary to say what shocks the impulses in Figure 3.18 are to and then which variable responses are to be summarized. Because impulse responses show the response of z_{t+M} to impulses in e_t, M (the horizon) needs to be set. The default below is for 10 periods. Lastly there are questions whether users want the results for impulses presented in terms of graphs or via a table. The question relating to standard errors will be returned to shortly.

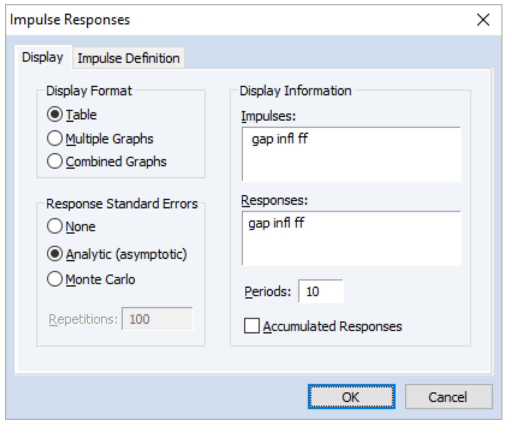


Figure 3.18: Generating Impulse Responses in EViews

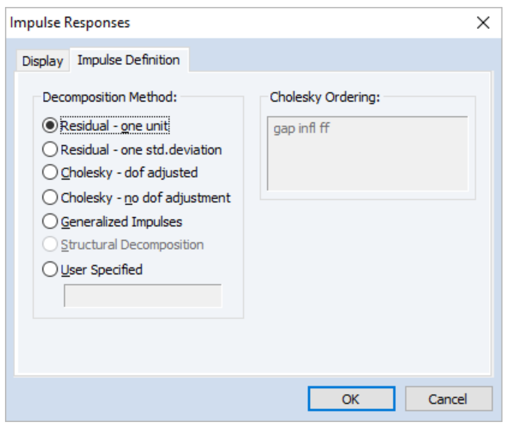


Figure 3.19: Types of Impulse Shocks in EViews

For the moment we will click on the ***Impulse Definition*** tab, bringing up Figure 3.19

If the first option is chosen then the impulse responses computed above are for one unit increases in the shocks coming from the fitted VAR. Choosing the Table option from Figure 3.18 and only asking for the impact of the shocks upon the output gap produces Table 3.19. For later reference we note that the two period ahead responses of the output gap to the three VAR errors are 1.205556, -0.109696 and 0.222451.

Often however responses are computed to one standard deviation shocks, i.e. the magnitude of the change in the $j'th$ shock would be $std(\varepsilon_{jt})$. The reason for this adjustment is that one unit may not be regarded as a shock of "typical" magnitude. If it is desired to have one standard deviation shocks then

Var: CHOMORENO Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Impulse Response to Nonfactorized One Unit Innovations

Response of GAP:

Period	GAP	INFL	FF
1	1.000000	0.000000	0.000000
	(0.00000)	(0.00000)	(0.00000)
2	1.216215	-0.058829	0.193684
	(0.12259)	(0.09654)	(0.10324)
3	1.205556	-0.109696	0.222415
	(0.19875)	(0.14889)	(0.14852)
4	1.109725	-0.112067	0.170645
	(0.23790)	(0.19151)	(0.16614)
5	0.992449	-0.120467	0.099129
	(0.25370)	(0.20721)	(0.16693)
6	0.869172	-0.131145	0.036128
	(0.26327)	(0.20306)	(0.16149)
7	0.749568	-0.145265	-0.013677
	(0.27045)	(0.18861)	(0.15292)
8	0.637292	-0.157941	-0.051147
	(0.27532)	(0.17066)	(0.14338)
9	0.534638	-0.167654	-0.078712
	(0.27673)	(0.15307)	(0.13354)
10	0.442370	-0.173440	-0.098345
	(0.27412)	(0.13750)	(0.12379)

Figure 3.20: Impulse Responses of the Output Gap, Inflation and the Intereest Rate to a Unit Change in the VAR Output Gap Equation Errors

the second option in Figure 3.19 would be selected. The last of the options is ***User Specified*** and this enables the investigator to set a number of shocks to specified magnitudes. The EViews manual has a good description of how to implement this option.

3.6 Standard Errors for Impulse Responses

Take the univariate AR(1) model again with impulse responses $D_l = b_{11}^l$. The estimated responses will be $\hat{D}_l = \hat{b}_{11}^l$, where $\hat{b}_{11}$ is the estimated AR(1) coefficient. The problem in attaching standard errors to $\hat{D}_l$ is clear from this expression. Even if $std(\hat{b}_{11})$ was known, $\hat{D}_l$ is formed from it in a non-linear way. There are two solutions to this. One is called "asymptotic" and utilizes what is known as the delta method. This says that, if $\hat{\psi} = g(\hat{\theta})$, where g is some function (e.g., $\theta = b_{11}$ and $g(\hat{\theta}) = \hat{b}_{11}^l$), then asymptotically the $var(\hat{\psi})$ can be approximated with

$$(\frac{\partial g}{\partial \theta}|_{\theta=\hat{\theta}})var(\hat{\theta})(\frac{\partial g}{\partial \theta}|_{\theta=\hat{\theta}}),$$

where $\frac{\partial g}{\partial \theta}|_{\theta=\hat{\theta}}$ means that the derivative is evaluated by setting $\theta = \hat{\theta}$. In the simple case being discussed this will be

$$(\frac{\partial g}{\partial b_{11}}|_{b_{11}=\hat{b}_{11}})var(\hat{b}_{11})(\frac{\partial g}{\partial b_{11}}|_{b_{11}=\hat{b}_{11}}),$$

and therefore $var(\hat{D}_l) = (l\hat{b}_{11}^{l-1})^2 var(\hat{b}_{11})$. There are matrix versions of this.

An alternative is to find the variance by bootstrap methods. These are more accurate than the asymptotic one if the normality assumption for $\hat{b}_{11}$ is incorrect. In this method it is assumed that the true value of b_{11} is $\hat{b}_{11}$ and then numbers are simulated from the AR(1). S simulations, $s = 1, .., S$, are performed. Each set of simulated data can be used to estimate the AR(1) and get an estimate $\tilde{b}_{11}^{(s)}$. In turn that produces an implied impulse response $\tilde{D}_l^{(s)}$. The mean and variance of $\hat{D}_l$ are then taken to be the first two sample moments of the $\tilde{D}_l^{(s)}, s = 1, .., S$. There are two types of bootstrap. The *parametric bootstrap* assumes a particular density function for the shocks, say $N(0, \sigma_1^2)$, and then uses a random number generator to get $\tilde{e}_{1t}^{(s)}$, where σ_1^2 is replaced by $\hat{\sigma}_1^2$ which is found from the data. The regular *bootstrap* uses the actual data residuals $\hat{e}_{it}$ as random numbers, re-sampling from these with a uniform random number generator to get $\tilde{e}_{1t}^{(s)}$. EViews gives the asymptotic (called ***Analytic*** in the screen shot above) and a simulation method (called ***Monte Carlo*** in the screen shot). Note that if the Monte Carlo option is chosen it is necessary to set the number of replications, i.e. S.

The Monte Carlo method in EViews is not the bootstrap method outlined above. In terms of the AR(1) it basically simulates different values of b_{11} by assuming that they come from a normal density with mean $\hat{b}_{11}$ and $var(\hat{b}_{11})$ that equals what was found from the data. Thus, for every new value of $b_{11}, \tilde{b}_{11}^{(s)}$, that is generated, it is possible to compute impulses $\tilde{D}_l^{(s)}$ and thereby get standard errors in the same way as was done with the bootstrap. Because the density of $\hat{b}_{11}$ is assumed to be normal the standard errors for impulse responses found in this way will only differ from the asymptotic results because the δ method uses a linear approximation to the $g(\theta)$ function. In most instances the differences between the two standard errors given by EViews will not be large.

Basically, the problems with standard errors for impulse responses come when $\hat{b}_{11}$ is *not normally distributed* in a finite sample. One case where this would be true is if b_{11} is close to unity, since the density function of $\hat{b}_{11}$ is then closer to the Dickey-Fuller density. There have been proposals to improve on the computation in this case, e.g. Kilian (1998) suggested the bootstrap plus bootstrap method, but none of these alternatives is in EViews. Therefore, it is always important to check how close to unity the roots of the VAR are. This can be done in EViews after estimation of a VAR by clicking on ***View***$\rightarrow$ ***Lag Structure***$\rightarrow$ ***AR Roots***. The information presented converts the estimated VAR to a companion form and then looks at the eigenvalues of the companion matrix. To see what this means take a VAR(2) in z_t. Then the companion form expresses the VAR(2) as a VAR(1) in the variables $w_t = \begin{bmatrix} z_t \\ z_{t-1} \end{bmatrix}$ and it has the form

$$w_t = \begin{bmatrix} z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ I_n & 0 \end{bmatrix} w_{t-1} + \begin{bmatrix} e_t \\ 0 \end{bmatrix}.$$

The eigenvalues being examined then are those of the matrix $\begin{bmatrix} B_1 & B_2 \\ I_n & 0 \end{bmatrix}$. For the US macro data the eigenvalues are less than 0.9 so that all methods of getting standard errors should work quite well. In fact the asymptotic and Monte Carlo methods give quite similar results for the small macro model.

3.7 Issues when Using the VAR as a Summative Model

We might distinguish three of these.

1. When there are missing variables from the VAR.
2. When there are latent (unobserved) variables not accounted for in the VAR.
3. When the relations are not linear. This can arise in a number of ways, e.g. in the presence of threshold effects or if there are categorical (dummy) variables in the VAR arising from latent or recurrent states.

3.7.1 Missing Variables

Theoretical models often have variables in them that are either not measured or it is felt that they are measured too imprecisely to be used in estimation, e.g. the capital stock of the macro economy. To this point the VAR has been described solely in terms of measured variables, so one needs to ask how the presence of missing or latent variables affects the nature of the VAR. In particular, what happens if the number of measured variables is less than the number in either a model that is being entertained or which one feels is needed to describe the macro economy?

To see what the effect of having missing variables is we take a simple example in which there should be two variables in the system being analyzed, z_{1t} and z_{2t}, but observations on only one of these, z_{1t}, is available. It will be assumed that the system in both variables is described by the restricted VAR(1) format

$$\begin{aligned} z_{1t} &= b_{11}^1 z_{1t-1} + b_{12}^1 z_{2t-1} + e_{1t} \\ z_{2t} &= b_{22}^1 z_{2t-1} + e_{2t}. \end{aligned}$$

Because only observations on z_{1t} are available it is necessary to see what the data generating process (DGP) of z_{1t} is.

Writing the equation for z_{2t} as $(1-b_{22}^1 L)z_{2t} = e_{2t}$, where L is the lag operator that lags z_t, i.e. $L^k z_t = z_{t-k}$, we have $z_{2t-1} = (1 - b_{22}^1 L)^{-1} e_{2t-1}$, and so

$$z_{1t} = b_{11}^1 z_{1t-1} + b_{12}^1 (1 - b_{22}^1 L)^{-1} e_{2t-1} + e_{1t}.$$

Hence

$$(1 - b_{22}^1 L) z_{1t} = b_{11}^1 (1 - b_{22}^1 L) z_{1t-1} + b_{12}^1 e_{2t-1} + (1 - b_{22}^1 L) e_{1t}.$$

It follows that

$$z_{1t} = (b_{11}^1 + b_{22}^1) z_{1t-1} - b_{11}^1 b_{22}^1 z_{1t-2} + b_{12}^1 e_{2t-1} + e_{1t} - b_{22}^1 e_{1t-1} \tag{3.16}$$

Equation (3.16) is an Autoregressive Moving Average (ARMA(2,1)) process. So the reduction in variables has changed the appropriate summative model for z_{1t} to an ARMA form from a VAR. This is a general result first noted by Wallis (1977) and Zellner and Palm (1974), i.e., if the complete set of variables has a VAR(p) for their DGP, then the DGP of the reduced set of variables will be a Vector Autoregressive Moving Average (VARMA) process. Of course a VARMA process can generally be thought of as a VAR(∞), and this leads to the possibility of findings that a high p is required may simply be reflecting the fact that not enough variables are present to capture the workings of the system. If this is so, then the solution is to change n and not p.

Variable reduction does not always result in a VARMA process. Suppose for example that $z_{2t} = 3z_{1t}$, i.e. the relation between the original and reduced set of variables is governed by an *identity.* Then

$$z_{1t} = (b_{11}^1 + 3b_{12}^1) z_{1t-1} + e_{1t}$$

which is still an AR(1). But, if the relation is $z_{2t} = 3z_{1t} + \phi z_{2t-1}$ (like with capital stock accumulation), we would have

$$\begin{aligned} z_{1t} &= b_{11}^1 z_{1t-1} + b_{12}^1 \frac{3z_{1t-1}}{(1-\phi L)} + e_{1t} \\ &\Longrightarrow \quad z_{1t} = (\phi_1 + b_{11}^1 + 3b_{12}^1) z_{1t-1} - \phi b_{11}^1 z_{1t-2} + e_{1t} - \phi e_{1t-1} \end{aligned}$$

which leads to an ARMA(2,1) process. The problem would also arise if $z_{2t} = 3z_{1t} + \eta_t$, where η_t is random, except that now the process for z_{1t} would be ARMA(1,1). So one needs to be careful in working with a reduced number of variables, and it is likely that it would be better to include the omitted variables in a VAR even if they are poorly measured. If this is impossible some proxy should be added to the VAR, e.g. investment should be present in it if the capital stock is omitted.

The effect can be quite large. Kapetanios *et al.* (2007) constructed a 26 endogenous variable model that was meant to emulate the Bank of England Quarterly model (BEQM) current around 2002-2008 (the smaller model had about half the number of variables of BEQM). Five impulse responses were constructed (corresponding to foreign demand, total domestic factor productivity, government expenditure, inflation and the sovereign risk premium) and then data on five variables were generated. These were standard variables in many open economy VARs - a GDP gap, inflation, a real interest rate, the real exchange rate and foreign demand. VARs were then fitted to the simulated data. Obviously there is a major reduction in the number of variables and it had a great impact on the ability to get the correct impulse responses. Indeed, it was found that a VAR(50) and some 30000 observations were needed to accurately estimate all impulse responses. It was found that the VAR orders determined by criteria such as AIC and SC were typically quite small (between four and seven) and led to major biases in the estimators of the impulse responses. Others have found similar results. A recent general discussion of the issues can be found in Pagan and Robinson (2016).

The implications of the discussion above would be that

- A small number of variables in a VAR probably means a high p is needed.
- If one finds a high p in a given VAR exercise this suggests that may be necessary to expand the number of variables rather than looking for a greater lag length.

In general it is important to think carefully about n, and good VAR modeling demands more than just the selection of p. Moreover, thought needs to be given to the nature of the variables included in the VAR, as well as to their number. As the simple analysis above showed, problems might be expected when stocks are omitted from the set of variables in the VAR. Indeed, this necessitated a much larger order VAR than is typically used in empirical exercises, where samples are relatively small. The neglect of stocks in most VAR work (generally the variables are just flows) is a potential problem that needs to be addressed.

Pagan and Robinson (2016) suggest that this is particularly so with VARs for small open economies featuring external assets. Apart from the implication that there may be a need for higher order VARs, the absence of stocks can lead to another problem that has been identified with VARs, viz. that of *non-invertible* VARMAs. In this instance the data requires a VARMA process that can not be captured by a VAR of *any* order. A simple example that shows this is the following, taken from Catao and Pagan (2011).

Suppose there is a desire to stabilize the level of debt relative to some target with a variable such as the primary deficit being manipulated to achieve that. If $\tilde{x}_t$ is the primary deficit and $\tilde{d}_t$ is the stock of debt defined as a gap relative to its desired equilibrium value, debt will accumulate as

$$\Delta\tilde{d}_t = \tilde{x}_t,$$

where we assume a zero real rate of interest for simplicity. In order to stabilize the debt we would have the primary deficit responding to debt levels and some activity variable $\tilde{y}_t$, such as an output gap

$$\tilde{x}_t = a\tilde{d}_{t-1} + c\tilde{y}_{t-1} + e_t, a < 0. \tag{3.17}$$

It will be assumed that $\tilde{y}_t$ is stationary with zero mean. Then

$$\Delta\tilde{d}_t = a\tilde{d}_{t-1} + c\tilde{y}_{t-1} + e_t, \tag{3.18}$$

and the debt gap converges to zero since $\tilde{y}_t$ is a stationary process.

Now suppose we attempted to use a VAR which did not include $\tilde{d}_t$, i.e. it only consisted of $\tilde{x}_t$ and $\tilde{y}_t$, as is common with many fiscal VARs and open economy studies. To see the effect of this it is necessary to solve for $\tilde{d}_t$ first and then substitute that variable out of the system. From (3.18)

$$\tilde{d}_t = (1-(1+a)L)^{-1}[c\tilde{y}_{t-1} + e_t],$$

so that the fiscal rule (3.17) can be expressed as

$$\tilde{x}_t = a(1-(1+a)L)^{-1}[c\tilde{y}_{t-2} + e_{t-1}] + c\tilde{y}_{t-1} + e_t. \tag{3.19}$$

Expanding (3.19) produces

$$\begin{aligned}\tilde{x}_t &= (1+a)\tilde{x}_{t-1} + ac\tilde{y}_{t-2} + ae_{t-1} + c\tilde{y}_{t-1} + e_t - (1+a)(c\tilde{y}_{t-2} + e_{t-1}) \\ &= (1+a)\tilde{x}_{t-1} + c\Delta\tilde{y}_{t-1} + \Delta e_t.\end{aligned}$$

The error term Δe_t is a non-invertible MA(1), meaning that there is no VAR representation for $\tilde{x}_t, \tilde{y}_t$ (there will be another equation for $\tilde{y}_t$ in the system but it may or may not involve the level of debt). Thus in this case compression of the set of variables results in a VARMA process but, importantly, one in which the MA term is not invertible. One of the first examples of a non-invertible process is given in Lippi and Reichlin (1994). It was an example of what has been described as *non-fundamentalness* and was first commented on by Hansen and Sargent (1991). There are many theoretical issues like this that we don't cover in the book and a good review of this topic is in Alessi *et al.* (2008). It might be observed that in the simple example we just presented the number of shocks is either less than or equal to the number of observables, so it is not true that the problem comes from having excess shocks (as was true of the first example we did where z_{2t} was omitted from the VAR when it should have been present).

3.7.2 Latent Variables

It often makes sense to account for unobserved variables more directly than just by increasing the order of the VAR. When it is believed that latent variables are present this is best handled using the *state space form (SSF)*

$$z_t^* = B_1 z_{t-1}^* + e_t,$$

where z_t^* are all the variables in the system but only some of them, z_t, are observed. We relate the observed variables to the total set z_t^* through the mapping

$$z_t = G z_t^*.$$

In most instances G will be a known selection matrix.

Because the likelihood is formulated in terms of z_t it is necessary to indicate how this is to be computed. Now the two equations describing z_t^* and z_t constitute a SSF, with the first equation being the *state dynamics* and the second equation being the *observation equation.* Then the likelihood depends on the conditional densities $f(z_t|Z_{t-1}), Z_{t-1} = \{z_{t-1}, z_{t-2}, .., z_1\}$. When the shocks e_t are normal, $f(z_t|Z_{t-1})$ is a normal density and so only $E(z_t|Z_{t-1}), var(z_t|Z_{t-1})$ need to be computed. The Kalman filter gives these via a recursive calculation. Consequently it is possible to set up a likelihood based on observables. So if one has latent variables in a VAR, the Kalman filter is a natural choice. EViews does perform Kalman Filtering and SSF estimation but, because it requires one to exit the VAR object, it is not possible to easily compute impulse responses etc. using pull-down menus.

3.7.3 Non-Linearities

3.7.3.1 Threshold VARs

The simplest non-linearity that has been proposed is the Vector STAR (VSTAR) model which has the form

$$z_t = B_1 z_{t-1} + F_1 z_{t-1} G(s_t, \gamma, c) + e_t$$

,where s_t is an observed threshold variable and c is the threshold value, i.e. $s_t > c$ shifts the transition function $G(\cdot)$ relative to what it was when $s_t \leq c$. In the STAR case G has the form

$$G(s_t, \gamma, c) = (1 + \exp(-\gamma(s_t - c)))^{-1}, \gamma > 0.$$

The nature of s_t varies but it is often a lagged value of z_t. One can have a higher order VAR than the VAR(1) above and s_t can be a vector.

Clearly the model is a non-linear VAR and, in the standard case where the same variables appear in every equation, it just involves performing non-linear rather than linear regression. There are some identification issues regarding γ and c, e.g. if $\gamma = 0$ then one cannot estimate c, but this is true for many non-linear regressions. One can think of this as a way of allowing for breaks in the parameters (for shifts in the density of z_t conditional upon z_{t-1} and s_t) not via dummy variables, but rather through a separate variable s_t. One difficulty which can arise occurs when s_t is a variable that is not included in the linear part of the model. In such a case if s_t actually did have a linear influence on z_t then it is very likely that $G(s_t)$ will seem to be significant, even when there is no non-linearity present. Hence at a minimum s_t should appear in the linear structure, which may mean treating it as an exogenous variable if it is not part of z_t. Recent applications (Hansen, 2011) of this model have been concerned with the impact of uncertainty upon the macro-economic system, leading to s_t being some volatility index such as the VIX.

There is no VSTAR option in EViews but it is available as an Add-On. This requires that one have R installed, since the code to perform the estimation is written in R.

3.7.3.2 Markov Switching process

The simplest Markov Switching (MS) model is

$$z_t = \delta_0 z_t^* + \delta_1(1 - z_t^*) + \sigma\varepsilon_t, \tag{3.20}$$

where z_t^* is a latent binary Markov process which is characterized by the transition probabilities $p_{ij} = \Pr(z_t^* = j | z_{t-1}^* = i)$. Here z_t^* takes values of 0 and 1 and p_{10} is the probability of going from a state of 0 in $t-1$ to 1 in t. It is possible to show that the MS structure implies that z_t^* is an AR(1) of the form

$$z_t^* = \phi_1 + \phi_2 z_{t-1}^* + v_t \tag{3.21}$$

$$var(v_t) = g(z_t^*) \tag{3.22}$$

The equations (3.20) and (3.21) look like those for an SSF but are different because $var(v_t)$ depends on the latent state z_{t-1}^* - in the SSF any time varying variance for v_t has to depend on *observable* variables. This feature implies that the density $f(z_t|Z_{t-1})$ is no longer normal, but it can still be computed recursively because it depends on only a finite number of states for z_t^*. A consequence of the non-normality in the density is that $E(z_t|Z_{t-1})$ is no longer a linear function of Z_{t-1} and so one ends up with a non-linearity in the VAR. Note that what is being modeled here is a shift in the parameters of the conditional density. The unconditional density parameters are constant, i.e. there are no breaks in the series. The mean, variance etc. of z_t are always the same.

One extension of the model above is to allow for more than a single variable in z_t. Krolzig and Toro (2004) treated European business cycle phases in this way, fitting a VAR(1) to GDP growth in six European countries from 1970Q3-1995Q4, and allowing for a three state MS process for the intercepts, i.e. z_t^* took three rather than two values. Algorithms exist to estimate such MS-VAR models (the original one being written by Krolzig) in Ox, Gauss and Matlab. These find $f(z_t|Z_{t-1})$ and thereby the likelihood. EViews 9 can estimate basic MS models, but not MS-VARs. When Krolzig and Toro estimated the MS-VAR model on the countries individually there was little evidence of a 3-state process, but it became much clearer when z_t included GDP growth from all six countries. There are a large number of parameters being estimated in their case - the VAR(1) in six series alone requires 36, while the MS(3) process makes for 12 (connected with μ_0 and μ_1) plus the 9 from the transition probabilities. In fact the number is greater, as the covariance matrix of the VAR shocks in the Krolzig and Toro application also shifts according to the states, making it a challenging estimation problem.

In the multivariate case and a single latent state z_t^* the equivalent of (3.20) would be

$$z_{jt} = \delta_{0j} z_t^* + \delta_{1j}(1 - z_t^*) + \sigma\varepsilon_{jt}, \tag{3.23}$$

and so

$$\bar{z}_t = \frac{1}{n}\sum_{j=1}^{n} z_{jt} = \frac{1}{n}\sum_{j=1}^{n}\delta_{1j} + (\frac{1}{n}\sum_{j=1}^{n}(\delta_{0j} - \delta_{1j}))z_t^* + \frac{\sigma}{n}\sum_{j=1}^{n}\varepsilon_{jt}.$$

Then, as $n \to \infty$, z_t^* would become a linear combination of the means $\bar{z}_t$ and it might be used to replace z_t^* in (3.23). Of course it is unlikely that most MS-VARs would have n being very large. It would also generally be the case that, if (3.23) was a VAR, then the solution for z_t^* would involve not only $\bar{z}_t$ but all the lagged values z_{jt-1}, and these could not be captured by $\bar{z}_{t-1}$.

Quite a lot of applications have been made with the MS-VAR model in recent times. One difficulty with the literature is that it mostly works with small VARs and that raises the issue of whether they are misspecified. It may be that the latent variable introduced into the VAR by the MS structure is just capturing the misspecification of the linear part. There is also an identification problem with MS models (the "labeling" problem). As Smith and Summers (2004, p2) say "These models are globally unidentified, since a re-labeling of the unobserved states and state dependent parameters results in an unchanged likelihood function". The labeling issue has been discussed a good deal in statistics - see Stephens (2000) - and a number of proposals have been made to deal with it, e.g. Frühwirth-Schnatter (2001), but few of these seem to have been applied to empirical work with MS models in economics.

3.7.3.3 Time Varying VARs

We mentioned earlier that there is an issue of breaks in the moments of the unconditional densities for variables in the VAR and there can also be shifts in the conditional variance. In terms of a scalar AR(1) $z_t = bz_{t-1} + \sigma\varepsilon_t$ the variance of z_t is $\frac{\sigma^2}{1-b}$, and it might be that both σ^2 and b shift in such a way that the variance is constant, and hence the unconditional moments do not have a break. What changes in this example is the parameters of the conditional density $f(z_t|z_{t-1})$. There is an emerging literature in which all the coefficients of the VAR are allowed to change according to a unit root process. Thus, if we form θ as a vector that represents the parameters that are in the matrix A_1, then the specification is

$$\theta_t = \theta_{t-1} + \nu_t.$$

The covariance matrix of v_t is either fixed or allowed to evolve as a stochastic volatility process. It is hard to give much of a sensible interpretation to the idea that θ evolves as a unit root process, but it makes more sense as a pragmatic device to get some feel for whether the VAR is stable enough to be useful for analysis. If one thinks of a simple model like $z_{1t} = b_t z_{1t-1} + e_{1t}$ with $b_t = b_{t-1} + \nu_t$ then, as the variance of ν_t becomes larger, the coefficient b_t wanders further from its initial value b_0. Basically the estimate for b_t will depend more on current data the larger is the variance of v_t i.e. one would down- weight older data when forming an estimate of "b". The stochastic volatility assumption for the variance of ν_t has a similar role, down-weighting "outliers" that could have too great an influence on the estimate of A_1. These algorithms are not yet implemented in EViews. Some Matlab routines are available. Mostly they use Bayesian techniques to estimate θ_t, although work for models like this suggests that one can estimate them with particle filter extensions of the Kalman filter. A recent contribution that seems to overcome the problem of θ_t being unbounded is Giratis *et al.* (2014) who write the model as $z_{1t} = b_{t-1}z_{1t-1} + e_{1t}$ and allow b_t to evolve as $b_t = b\frac{\phi_t}{\max_{0\leq k\leq t}|\phi_k|}$ $(b \in (0,1))$, where $\phi_t = \phi_0 + \sum_{j=1}^{t} v_t$. This enables them to handle shifts via kernel based smoothing methods.

3.7.3.4 Categorical Variables for Recurrent States

Categorical variables S_t are sometimes added into VARs to represent recurrent states. These are generally constructed from some observed variable. For example we might have $S_t = 1(\Delta y_t > 0)$, where $1(\cdot)$ is the indicator function taking the value unity if the event in brackets is true, and zero if it is false. Here Δy_t might be an observed growth rate. The non-linearity comes from the indicator function. Adding indicators of recurrent events, S_t, into VARs can raise complex issues. Many of these stem from the fact that the S_t are constructed from some variable like y_t and this will influence their nature. In particular they show a non-linear dependence on the past and possibly the future.

To see this non-linear dependence take the business cycle indicators constructed by the NBER. Because the NBER insist that recessions must last two quarters the S_t will evolve as a Markov Chain of at least second order. To see this take data on the NBER S_t over 1959/1 to 1995/2 (S_t equals one for an expansion and zero for a contraction) and fit a second order Markov chain to S_t yielding[8]

$$S_t = \underset{(3.8)}{0.4} + \underset{(5.6)}{0.6S_{t-1}} - \underset{(-3.8)}{0.4S_{t-2}} + \underset{(3.1)}{0.35S_{t-1}S_{t-2}} + \eta_t. \tag{3.24}$$

The non-linear term that comes with a second order Markov Chain ($S_{t-1}S_{t-2}$) is clearly important.

There are further difficulties in using S_t in a VAR as it may depend on future values of z_t. This occurs since in order to define the value of S_t it is necessary to know at what point in time peaks and troughs in activity occur. When S_t are NBER business cycle states then for a peak to occur at t it is necessary that Δz_{t+1} and $\Delta_2 z_{t+2}$ both be negative - see Harding and Pagan (2002). Consequently S_{t-1} depends on Δz_{t+1} and $\Delta_2 z_{t+2}$ and therefore cannot be treated as predetermined, meaning that a regression of z_t against S_{t-1} would yield inconsistent estimators of any VAR coefficients. There is a recent literature that replaces S_t with a continuous variable $\Phi(\Delta y_{t-1}, \Delta_2 y_{t-1})$, where Φ is a function that lies between zero and one. Φ_t is now predetermined but it is clear that it will lag behind the point in time that the NBER-defined recessions and expansions start. Because of these complications with S_t one cannot use existing ways of estimating VARs that are available in EViews. Harding and Pagan (2011) contain a detailed discussion of the problems that arise.

[8]Newey-West HAC t-ratios in brackets using a window-width of four periods.

Chapter 4

Structural Vector Autoregressions with I(0) Processes

4.1 Introduction

The errors in a VAR will generally be correlated. For this reason it is hard to know how to use an impulse response function, as these are meant to measure the change in a shock *ceteris paribus.* If the shocks are correlated however, one can't hold other shocks constant when a shock occurs. In the next section two methods are outlined that combine the VAR errors together so as to produce a set of uncorrelated shocks for which impulse responses can be computed – these are named according to the mathematical method employed to obtain the uncorrelated shocks. In the following section Generalized Impulse Responses (GIR) are discussed. This method does not re-define the shocks as above but instead computes impulse responses to the VAR errors which make some allowance for the fact that they are correlated. In the remainder of the chapter the strategy of finding uncorrelated shocks is resumed. However attention switches to providing economic justifications for the processes that lead to shocks that are uncorrelated.

4.2 Mathematical Approaches to Finding Uncorrelated Shocks

For convenience we will begin by working with a VAR(1)

$$z_t = B_1 z_{t-1} + e_t,$$

where $E(e_t) = 0$ and $cov(e_t) = \Omega_R$. Consider combining the errors e_t together with a non-singular matrix P so as to produce a set of uncorrelated shocks v_t, i.e. $e_t = Pv_t$ and the ν_{it} are uncorrelated with each other. Thus

$$z_t = B_1 z_{t-1} + Pv_t,$$

and the response of z_t to v_t will be P. Accordingly, the problem is to find a P matrix such that ν_t is uncorrelated, i.e. $cov(\nu_t) = F$, where F is a diagonal matrix. There are two approaches to finding such a P :

1. The singular value decomposition (SVD) of the matrix Ω_R is $\Omega_R = UFU'$, where $U'U = I, UU' = I$ and F is a diagonal matrix. Therefore setting $P = U$ will work, as $cov(e_t) = \Omega_R = PFP'$. Consequently, contemporaneous impulse responses to the shocks ν_t would be P.

2. The Cholesky decomposition is $\Omega_R = A'A$ where A is a triangular matrix. Hence setting $P = A', F = I$ also works, although in this case the variances of the ν_t are unity. But if one prefers the variances of shocks not to be unity, this can be accounted for by allowing the diagonal elements of A to capture the standard deviations. The contemporaneous impulse responses to unit shocks will be given by A'.

Because the orthogonal shocks will be different when using these different methods, so to will be the impulse responses. But there is no way of choosing between the two approaches as they both replicate Ω_R.

To perform a Cholesky decomposition in EViews take the small macro model with the variables *gap, infl, ff*. Following the instructions to Figure 2.4 click on ***Impulse***$\rightarrow$ ***Impulse definition*** $\rightarrow$***Cholesky - dof adjusted***. This produces the impulse responses for those shocks. When computing the Cholesky decomposition in EViews note that one has to describe the order that the variables enter into the VAR. In this case they are ordered as entered, namely *gap, infl, ff*.

This brings up a key difference between the Cholesky and SVD approaches. When P is formed from the SVD it orders the (orthogonal) variates from most variation to the least. This ranking does not change if the ordering of the original variables in the VAR is changed, i.e. if the variables were entered into EViews as *ff, gap, infl* rather than *gap, infl, ff*. Clearly, the Cholesky decomposition results will change but not those for the SVD.

4.3 Generalized Impulse Responses

A different approach is not to construct new shocks ν_t but to investigate the impact on the variables z_{jt} of changes in e_{jt}. Because the errors are correlated it needs to be recognized that the impact of any change in an error cannot be found directly, but has to allow for the fact that changes made to e_{jt} will also mean changes in e_{kt}. Consequently, the final impact of a change in e_{jt} on z_{lt} needs to take this into account. To see how this is done is a simple context take the two variable VAR

$$z_{1t} = b_{11}^1 z_{1t-1} + b_{12}^1 z_{2t-1} + e_{1t}$$
$$z_{2t} = b_{21}^1 z_{1t-1} + b_{22}^1 z_{2t-1} + e_{2t},$$

where $\Omega_R = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}$. Now consider increasing e_{1t} by one standard deviation $\sqrt{\omega_{11}}$. What is the effect on z_2? To answer that we look at the MA form for the VAR, which will be $z_{2t} = e_{2t} + D_{21}^1 e_{1t-1} + D_{22}^1 e_{2t-1} + ...$ Now, if it is assumed that e_t follows a bivariate normal, then $e_{2t} = \frac{\omega_{21}}{\omega_{11}} e_{1t} + \eta_{2t}$, where η_{2t} is uncorrelated with e_{1t}, i.e. η_{2t} remains unchanged as e_{1t} is varied. From this, the effect on z_{2t} of a change in $e_{1t}$of magnitude $\sqrt{\omega_{11}}$ will be $\frac{\omega_{21}}{\omega_{11}}\sqrt{\omega_{11}}$. Basically this is equal to $E(z_{2t}|e_{1t} = \sqrt{\omega_{11}}) - E(z_{2t}|e_{1t} = 0)$.

Continuing along these same lines, but for longer horizon impulse responses, look at z_{2t+1}. Then there will be a *direct effect* due to e_{1t} changing and an *indirect* effect due to the change in e_{1t} affecting e_{2t}. Hence the *total* effect of a change in e_{1t} of $\sqrt{\omega_{11}}$ on z_{2t+1} will be $D_{21}^1\sqrt{\omega_{11}} + D_{22}^1 \frac{\omega_{21}}{\omega_{11}}\sqrt{\omega_{11}}$.

This method was originally due to Evans and Wells (1983, 1986) but has been popularized under the title of generalized impulse responses (GIRs) by Pesaran and Shin (1998). It should be noted that computation of a generalized impulse response function for e_{jt} can be done by placing z_{jt} first in the ordering of variables, and then calculating the impulse response using the Cholesky decomposition. This shows that the ordering of variables does not matter when computing GIRs, as each variable z_{jt} takes its turn at the top of the order to define a shock, after which the Cholesky decomposition gives the impulse responses to this shock. It is necessary to order the variables n times to get all the GIRs.

What is the use of these GRIs ? First, there are no names for the shocks being applied. One is just combining the VAR errors, so the only names they have are the first, second etc. equation shocks, and that does not seem particularly attractive. Second, each of the shocks comes from a different recursive model, not a single model. It has been argued that GIRs are useful for studying the persistence of shocks ("persistence profiles"). But persistence just depends on the eigenvalues of B_1, and these are easy to find from EViews pull-down menus, as explained in the previous chapter. Consequently, it is hard to see the value in doing a GI analysis.

4.4 Structural VAR's and Uncorrelated Shocks: Representation and Estimation

4.4.1 Representation

The more standard approach is to note that the correlations between shocks arise due to contemporaneous correlations between variables and so, instead of having a variable depending only upon past values of other variables, one needs to look at systems where each variable can also depend on the contemporaneous values of other variables. Then, because one has (hopefully) captured the contemporaneous effects, the errors in the structural equations can now be taken to be uncorrelated. Creating this new system will result in it having a different function. It now performs an *interpretative* task. In more familiar terms it consists of structural (simultaneous) equations rather than a reduced form like the VAR. Because structural equations originally derived from the idea that they reflected decisions by agents this approach can be said to have economic content.

To be more specific, the resulting structural VAR (SVAR) system of order p will be:

$$A_0 z_t = A_1 z_{t-1} + ... + A_p z_{t-p} + \varepsilon_t,$$

where now the shocks ε_t are taken to be uncorrelated, i.e. $E(\varepsilon_t) = 0, cov(\varepsilon_t) = \Omega_S$ and Ω_S is a diagonal matrix. The elements in the matrices will follow the same conventions as previously for the lagged ones, i.e. the $(i,k)^{th}$elements in the j'^{th} lag matrix will be $\{a^j_{ik}\}$. It is necessary to look more carefully at A_0. It will be defined as

$$A_0 = \begin{bmatrix} a^0_{11} & -a^0_{12} & . & . \\ -a^0_{21} & a^0_{22} & -a^0_{23} & . \\ . & & & \\ . & & & \end{bmatrix},$$

where the signs on a^0_{ij} are chosen so as to enable each of the equations to be written in regression form.

To explore this more carefully take the following "market model" (which could be the demand and supply for money in terms of interest rates)

$$q_t - a^0_{12} p_t = a^1_{11} q_{t-1} + a^1_{12} p_{t-1} + \varepsilon_{1t} \quad (4.1)$$
$$p_t - a^0_{21} q_t = a^1_{21} q_{t-1} + a^1_{22} p_{t-1} + \varepsilon_{2t} \quad (4.2)$$
$$var(\varepsilon_{1t}) = \sigma^2_1, var(\varepsilon_{2t}) = \sigma^2_2, cov(\varepsilon_{1t}\varepsilon_{2t}) = 0,$$

where q_t is quantity and p_t is price.

This reduces to a VAR of the form

$$\begin{aligned} q_t &= b^1_{11} q_{t-1} + b^1_{12} p_{t-1} + e_{1t} \\ p_t &= b^1_{21} q_{t-1} + b^1_{22} p_{t-1} + e_{2t}. \end{aligned}$$

Now the model (4.1) - (4.2) implies that $A_0 = \begin{bmatrix} 1 & -a^0_{12} \\ -a^0_{21} & 1 \end{bmatrix}$ and it is said to be in *normalized* form, i.e. every equation has a "dependent variable" and every shock ε_{it} has a variance of σ^2_i. In contrast the *unnormalized form* would be

$$\begin{aligned} a^0_{11} q_t - a^0_{12} p_t &= a^1_{11} q_{t-1} + a^1_{12} p_{t-1} + \eta_{1t} \\ a^0_{22} p_t - a^0_{21} q_t &= a^1_{21} q_{t-1} + a^1_{22} p_{t-1} + \eta_{2t} \\ var(\eta_{1t}) &= 1, var(\eta_{2t}) = 1, cov(\eta_{1t}\eta_{2t}) = 0. \end{aligned}$$

In this latter form A_0 is left free and it can be assumed that the variances of the η_{it} are unity, since a_{ii}^0 is effectively accounting for them. By our definitions we would have $\varepsilon_{jt} = \sigma_j \eta_{jt}$, where σ_j is the standard deviation of ε_{jt}. Because ε_{jt} is just a re-scaled version of η_{jt} the properties of one representation apply to the other.

More generally a SVAR could be written as $A_0 z_t = A_1 z_{t-1} + B\eta_t$, with $var(\eta_{it})$ set to unity and with A_0 and B being chosen to capture the contemporaneous interactions among the z_t, along with the standard deviations of the shocks. This is actually the way EViews writes an SVAR, and the representation makes it possible to shift between whether the system is normalized or unnormalized depending on how one specifies $A = A_0$ and B. In EViews what we call η_t is labeled as u_t. Throughout the monograph whenever we want shocks that have a unit variance we will use η_t to mean this.

4.4.2 Estimation

Now as readers are probably aware there is an *identification problem* with simultaneous equations, namely it is not possible to estimate all the coefficients in $A_0, A_1, ..., A_p$ without some restrictions. However, compared to the standard simultaneous equations set-up described in Chapter 1, there is now an extra set of constraints in that the shocks ε_t are uncorrelated.

The summative model is meant to contain all the information present in the data in a compact form. In order to illustrate the issues in moving to an SVAR(1) let us assume that a VAR(1)

$$z_t = B_1 z_{t-1} + e_t$$

is the summative model. Putting $n = 2[3]$ (*the square brackets [·] will show the* $n = 3$ *case*) will help fix the ideas.[1] Then the summative (VAR) model has n^2 $(= 4[9])$ elements in B_1 and $\frac{n(n+1)}{2}(= 3[6])$ elements in the covariance matrix of e_t (symmetry means there are not n^2 values). The SVAR(1) coefficients have to be estimated somehow from these $n^2 + \frac{n(n+1)}{2}$ (=7[15]) pieces of information. Turning to the SVAR there are $n^2 - n$ (=2[6]) elements in A_0 (after normalization), n (=2[3]) variances of the shocks, and n^2 (=4[9]) unknowns in A_1, giving a total of $2n^2(= 8[18])$ parameters to estimate. So the number of parameters to be estimated in the SVAR exceeds that in the VAR. Consequently, it is not possible to recover all the coefficients in the SVAR(1) from the VAR(1). An extra $2n^2 - n^2 + \frac{n(n+1)}{2} = \frac{n(n-1)}{2}(= 1[3])$ restrictions are needed on A_0 and/or A_1. Finding such restrictions is a challenge and will be the subject of the remainder of this chapter and later ones. For now it is useful to assume that they have been found and to ask how the resulting SVAR would be estimated and how impulse responses would be formed after the estimation.

[1] $n = 3$ is of interest since the small macro model worked with in Chapter 2 had three variables.

4.4.2.1 Maximum Likelihood Estimation

The approximate log likelihood for an SVAR(p) is[2]

$$L(\theta) = cnst + \frac{T-p+1}{2}\{\ln |A_0|^2 + \ln |\Omega_S^{-1}|\}$$
$$-\frac{1}{2}\sum_{t=p+1}^{T}(A_0 z_t - A_1 z_{t-1} - .. - A_p z_{t-p})'\Omega_S^{-1}(A_0 z_t - A_1 z_{t-1} - .. - A_p z_{t-p}).$$

The term involving A_0 clearly involves parameters that need to be estimated. However, before EViews 8 it was not possible for a user to write a program that would maximize this likelihood, since a determinant like $|A_0|$ could not depend on unknown parameters. EViews 8 introduced the *optimize*() command that removed this constraint. The standard FIML estimator in EViews might have been considered as a potential estimator, but until EViews 9.5 it was not possible to to constrain the the structure of the structural error covariance matrix to be diagonal. The option to do this was introduced in EViews 9.5, and in what follows we make use of both *optimize*() and the enhanced FIML estimator to estimate the SVAR models considered. The most common assumption across the models is that the covariance matrix of structural equation shocks is diagonal.

4.4.2.2 Instrumental Variable (IV) Estimation

Often we will focus upon the estimation of SVAR systems using instrumental variables (IV) rather than MLE. There are some conceptual advantages to doing so, which will become apparent in later chapters that consider complex cases. Therefore it is worth briefly mentioning this procedure and some complications that can emerge when using it. To this end consider the single equation

$$y_t = w_t\theta + v_t,$$

where w_t is a stochastic random variable such that $E(w_t v_t) \neq 0$. Then application of OLS would give biased estimates of θ. However, if it is possible to find an instrument for w_t, x_t, such that $E(x_t v_t) = 0$, the IV estimator of θ would then be defined as

$$\hat{\theta} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} w_t x_t}.$$

Rather loosely we can say that $\hat{\theta}$ will be a consistent estimator of θ provided that

1. The instrument is correlated with w_t (the relevance condition).
2. It is uncorrelated with v_t (the validity condition).

[2] Remember that the exact log likelihood requires the unconditional density of $z_1,...,z_p$ as well. Throughout the rest of the book we will ignore the distinction and simply refer to the approximate log likelihood as the log likelihood and the estimator maximizing it as the MLE.

Instruments are said to be *weak* if the correlation of w_t with x_t is low. It is hard to be precise about this in an empirical context but, in the case of a single instrument, a correlation that is less than 0.1 would probably be considered as weak. If there is more than one instrument available, i.e. x_t is a vector, and we want to know if x_{1t} is a weak instrument, then empirical researchers often look at the F test that the coefficient of x_{1t} in the regression of w_t on x_t is zero. A value of less than 10 would be the equivalent indicator to the correlation above. When there are weak instruments $\hat{\theta}$ is generally not normally distributed, even in large samples, and often has a finite sample bias. These facts mean that it is hard to do inferences about the value of θ. As we go along, a number of examples where weak instruments can arise in SVARs will be encountered.

One method that is often suggested as allowing inferences that are robust to weak instruments is that of the Anderson-Rubin (1949) test. Suppose we write the equation above in matrix form as $y = W\theta + v$ with the IV estimator being $\hat{\theta} = (X'W)^{-1}X'y$. Then if we wish to test that $\theta{=}\theta^*$ we could do this by testing if $E\{X'(y-W\theta^*)\}{=}0$ using $\{X'(y-W\theta^*)\}$. This is a standard method of moments test and the variance of this is well defined, giving a test statistic that will be asymptotically χ^2. Now

$$X'(y - W\theta^*) = X'(y - W\hat{\theta} + W(\hat{\theta} - \theta^*)),$$

and this equals $X'W(\hat{\theta} - \theta^*)$ because $X'(y - W\hat{\theta}) = 0$ by the definition of the IV estimator. We can see from this that the distribution of $(\hat{\theta} - \theta^*)$ can be badly behaved whenever $X'W$ is a random variable with mean close to zero (as happens with weak instruments). Thus the advantage of the AR test is that it avoids working with $\hat{\theta}$. In practice θ^* is varied and confidence intervals are found. EViews does not provide the AR test although code could be written for it.

Now suppose that the SVAR equations were estimated via instrumental variables rather than MLE. If the SVAR was *exactly identified* (which mostly they are) then the MLE and the IV estimators are identical. This was proved in the context of the simultaneous equations literature by Durbin (1954) and Hausman (1975). Consequently the choice of which method is to be used in the exactly identified case must reside in computational and pedagogical considerations. One advantage of the IV approach is that it can point out cases where weak instruments arise. Although these must equally affect the MLE it is often not so obvious. Another advantage of the IV methodology is that it easily handles over-identified models in EViews.

4.5 Impulse Responses for an SVAR: Their Construction and Use

4.5.1 Construction

The impulse responses to VAR shocks were found from the MA form $z_t = D(L)e_t$. Impulse responses to structural shocks follow by using the relation between the VAR and SVAR shocks of $e_t = A_0^{-1}\eta_t$, where the use of η_t rather than ε_t points to an un-normalized form, i.e. the standard deviations of the shocks are absorbed into the diagonal elements of A_0. The MA representation for a VAR was given in Chapter 3 as $z_t = D(L)e_t$, leading to $z_t = D(L)A_0^{-1}\eta_t = C(L)\eta_t$ as the MA form for the SVAR. Thus $C(L) = D(L)A_0^{-1}$. From Chapter 3 the D_j are generated recursively as

$$\begin{aligned}
D_0 &= I_n \\
D_1 &= B_1 D_0 \\
D_2 &= B_1 D_1 + B_2 D_0 \\
&. \\
&. \\
D_j &= B_1 D_{j-1} + B_2 D_{j-2} + ... + B_p D_{j-p}.
\end{aligned}$$

Accordingly, equating terms in L from $C(L) = D(L)A_0^{-1}$ means that

$$\begin{aligned}
C_0 &= A_0^{-1} && (4.3) \\
C_1 &= D_1 A_0^{-1} = B_1 D_0 A_0^{-1} = B_1 C_0 && (4.4) \\
C_2 &= D_2 A_0^{-1} = (B_1 D_1 + B_2 D_0)A_0^{-1} = B_1 C_1 + B_2 C_0 \\
&. \\
&. \\
C_j &= D_j A_0^{-1} = (B_1 D_{j-1} + ... + B_p D_{j-p})A_0^{-1}. && (4.5)
\end{aligned}$$

From this it is clear that, when $j \geq p$, C_j can be generated recursively using (4.5) as

$$C_j = B_1 C_{j-1} + ... + B_p C_{j-p},$$

with the initial conditions $C_0, .., C_{p-1}$ being found from (4.3), (4.4) etc. Now, because the D_j can be computed by knowing just the VAR coefficients $B_1 ... B_p$, they do not depend in any way upon the structure of the model. Hence, once a structure is proposed that determines C_0, all the C_j can be found, emphasizing that the key issue for structural impulse responses is how C_0 is to be estimated.

4.5.2 Variance and Variable Decompositions

Because the shocks ε_t (or η_t) have been found, questions naturally arise about the importance of one shock versus others in explaining z_t. Two methods of using the impulse responses to answer such questions have emerged. One of these decomposes the variances of the forecast errors for z_{jt+h} using information at time t into the percentage explained by each of the shocks. The other gives a dissection of the variables z_{jt} at time t according to current shocks (and their past history).

Suppose that some information is available at time t and it is desired to predict z_{t+2} using a VAR(2). Then

$$\begin{aligned} z_{t+2} &= B_1 z_{t+1} + B_2 z_t + e_{t+2} \\ &= B_1(B_1 z_t + B_2 z_{t-1} + e_{t+1}) + B_2 z_t + e_{t+2} \\ &= (B_1^2 + B_2) z_t + B_1 B_2 z_{t-1} + B_1 e_{t+1} + e_{t+2}. \end{aligned}$$

Since z_t, z_{t-1} are known at time t the 2-step prediction error using information at that time will be $B_1 e_{t+1} + e_{t+2}$. Because $C_1 C_0^{-1} = B_1$ and $C_0 = A_0^{-1}$ from (4.3) the prediction errors can be re-expressed as

$$\begin{aligned} B_1 e_{t+1} + e_{t+2} &= C_1 A_0 e_{t+1} + e_{t+2} \\ &= C_1 A_0 A_0^{-1} \eta_{t+1} + A_0^{-1} \eta_{t+2} \\ &= C_1 \eta_{t+1} + C_0 \eta_{t+2}. \end{aligned}$$

It therefore follows that the variance of the two-step ahead prediction errors is

$$\begin{aligned} V_2 &= var(C_0 \eta_{t+2}) + 2cov(C_0 \eta_{t+2}, C_1 \eta_{t+1}) + var(C_1 \eta_{t+1}) \\ &= C_0 C_0^{'} + C_1 C_1^{'}, \end{aligned}$$

since $cov(\eta_t) = I_2$. Taking $n = 2$ and partitioning the matrices as

$$C_0 = \begin{bmatrix} c_{11}^0 & c_{12}^0 \\ c_{21}^0 & c_{22}^0 \end{bmatrix}, C_1 = \begin{bmatrix} c_{11}^1 & c_{12}^1 \\ c_{21}^1 & c_{22}^1 \end{bmatrix},$$

the variance of the two-step prediction error for the first variable will be

$$\Delta = (c_{11}^0)^2 + (c_{12}^0)^2 + (c_{11}^1)^2 + (c_{12}^1)^2.$$

Hence the first shock contributes $(c_{11}^0)^2 + (c_{11}^1)^2$ to the variance of the prediction error of z_{1t}, meaning that the fraction of the 2-step forecast variance accounted for by it will be $\frac{(c_{11}^0)^2 + (c_{11}^1)^2}{\Delta}$. The Forecast Error Variance Decomposition (FEVD) gives these ratios for forecasts made at t into the future but expressed as percentages

This information is available from EViews. After fitting an SVAR and using the Cholesky option when ***Impulse*** is chosen, one then selects ***View*** $\rightarrow$ ***Variance Decomposition*,** filling in the window that is presented. Using the ordering of the variables as *gap*, *infl*, *ff* the percentage of the variance of the ten periods ahead forecast error for inflation explained by the first orthogonal shock in the system is 15.92%, by the second one 75.85%, and the last shock explains 8.24%. Exactly why we are interested in looking at what shocks explain the forecast variance is unclear (except of course in a forecasting context). It is sometimes argued that this decomposition provides information about business cycle causes but, as pointed out in Pagan and Robinson (2014), its connection with business cycles is very weak.

The fundamental relation used to get the variance decomposition was

$$z_{t+2} = (B_1^2 + B_2)z_t + B_1 B_2 z_{t-1} + C_1 \eta_{t+1} + C_0 \eta_{t+2}.$$

A different viewpoint regarding the influence of shocks is therefore available by observing that, after starting with some initial values for z_t $(t \geq p)$, z_t can be expressed as a function of the standardized shocks $\{\eta_{t-j}\}_{j=0}^{t}$ weighted by the impulse responses. This is a useful decomposition since it shows what shocks are driving the variables z_t over time. Unfortunately, this variable decomposition is not directly available in EViews although an add-in (user-program) for historical decomposition (*hdecomp*) is available for download from www.eviews.com .

4.6 Restrictions on an SVAR

Finding enough restrictions on the SVAR so that it is identified is a challenge. Essentially it involves telling a story about how the macro economy works by stating conditions that enable the differentiation of the shocks. In this section three types of restrictions are used.

1. Making the system recursive.
2. Imposing parametric restrictions on the A_0 matrix.
3. Imposing parametric restrictions on the impulse responses to the shocks ε_t.

In the next chapter the restrictions are expanded to using sign restrictions on the impulse responses to the shocks ε_t to differentiate between them. Lastly, Chapters 6 and 7 look at using the long run responses that variables have to shocks as a way of discriminating between them.

4.6.1 Recursive Systems

The simplest solution to identification is to make the system *recursive.* As mentioned in Chapter 1 this assumes that A_0 is (typically) lower triangular and the structural shocks are uncorrelated. It was originally proposed by Wold (1951) as a method of identifying the parameters of structural equations. Wold's suggestion reduces the number of unknown parameters to exactly the number estimated in the summative model. The combination of triangularity and uncorrelated shocks means that a *numerical method* for estimating a recursive system is the Cholesky decomposition, and so this gives an economic interpretation of what the latter does. Basically it is a story about a given endogenous variable being determined by those "higher up" in the system but not those "lower down".

It is recommended that this solution be considered first and then ask if there is something unreasonable about it. If there is, then ask how the system should be modified? Because of its connection with the Cholesky decomposition it is the case that a recursive system will have an *ordering* of variables, but now theoretical or institutional ideas should guide the ordering choice. In this monograph the ordering is z_{1t}, z_{2t}, z_{3t} etc., i.e.A_0 is lower triangular, but sometimes you will see researchers make A_0 upper triangular, and then the ordering is from the bottom rather than the top.

As a simple example take the market model, where a recursive system could be

$$\begin{aligned} q_t &= a_{11}^1 q_{t-1} + a_{12}^1 p_{t-1} + \varepsilon_{S,t} \qquad (4.6) \\ p_t - a_{21}^0 q_t &= a_{21}^1 q_{t-1} + a_{22}^1 p_{t-1} + \varepsilon_{D,t}. \qquad (4.7) \end{aligned}$$

The idea behind this system is that quantity supplied does not depend contemporaneously on price, and that could be justified by institutional features. It seems reasonable to assume that the demand and supply shocks are uncorrelated since the specification of the system allows prices and quantities to be correlated. Of course if there was some common variable that affected both quantity and price, such as weather, unless it is included in each curve the structural errors in both equations would incorporate this common effect, and so the structural errors could not be assumed to be uncorrelated. This underscores the importance of making n large enough.

As mentioned in Chapter 1 it seems clear that the applicability of recursive systems will depend on the length of the observation period. If all one has is yearly data then it is much harder to come up with a plausible recursive system. In contrast, with daily data it is very likely that systems will be recursive. An alternative to the recursive system in (4.6) - (4.7) would be that price is determined (ordered) before quantity. The two systems are *observationally equivalent* since they replicate the estimated VAR and cannot be separated by any test using the data. Some other criterion is needed to favor one over the other. This might be based on institutional knowledge, e.g. that the quantity of fish might be fixed in a market and, if storage is difficult, price has to clear the market.

Each of the systems above solves the *structural identification* problem, reducing the number of parameters to be estimated to seven, namely:

$$(a_{21}^0, a_{ij}^k, var(\varepsilon_S), var(\varepsilon_D)).$$

MLE can be used to estimate the system, and in this case OLS is exactly identical to MLE.[3] To see this observe that (4.6) can clearly be estimated by OLS, and this is also true of (4.7), since $E(q_t \varepsilon_{Dt}) = E(a_{11}^1 q_{t-1}\varepsilon_{Dt} + a_{12}^1 p_{t-1}\varepsilon_{Dt} + \varepsilon_{S,t}\varepsilon_{Dt}) = 0$ due to the structural shocks being uncorrelated. An alternative way to estimate (4.7) that will be used later is to take the residual from the first equation and use it as an instrument for q_t in the second equation. *All of these approaches are identical* for a recursive model that is exactly identified. If there

[3]Because $|A_0| = 1$ in recursive (normalized) systems it is easily seen that MLE and OLS are identical.

were less (more) parameters in $\{A_0, A_1, cov(\varepsilon_t)\}$ compared to $\{B_1, cov(e_t))\}$ then there will be *over-(under-)identification,* and the estimators can differ.

4.6.1.1 A Recursive SVAR with the US Macro Data

We will now return to the three variable US macro model of the previous chapters. For simplicity of exposition it will be assumed to have the recursive SVAR(1) form. Later it will be implemented as an SVAR(2).

$$\begin{aligned} y_t &= a_{11}^1 y_{t-1} + a_{12}^1 \pi_{t-1} + a_{13}^1 i_{t-1} + \varepsilon_{1t} & (4.8) \\ \pi_t &= a_{21}^0 y_t + a_{21}^1 y_{t-1} + a_{22}^1 \pi_{t-1} + a_{23}^1 i_{t-1} + \varepsilon_{2t} & (4.9) \\ i_t &= a_{31}^0 y_t + a_{32}^0 \pi_t + a_{31}^1 y_{t-1} + a_{32}^1 \pi_{t-1} + a_{33}^1 i_{t-1} + \varepsilon_{3t}. & (4.10) \end{aligned}$$

Equations (4.8)-(4.10) provide a recursive story about "inertial responses" since they are is based on

1. Interest rates having no effect on the output gap for one period
2. There is no *direct* effect of current interest rates upon inflation.
3. There is an interest rate rule in which the monetary authority responds to the current output gap and inflation.

The shocks in this system are given the names of demand (ε_{1t}), supply/costs (ε_{2t}) and monetary/ interest rate (ε_{3t}). Effectively, the story reflects institutional knowledge about rigidities and caution in the use of monetary policy.

4.6.1.2 Estimating the Recursive Small Macro Model with EViews

The data used is that of Chapter 2. We will estimate the SVAR in (4.8) - (4.10) using EViews but with 2 lags. The simplest way to fit a recursive model is to just utilize a Cholesky decomposition after the VAR has been estimated. Estimation and the derivation of impulse responses from a Cholesky decomposition was described in Section 4.2, so these will be the impulse responses for the recursive system in (4.8) - (4.10). Figure 4.1 graphs these.

There are other ways of estimating the SVAR that will be used extensively in what follows. It is useful therefore to illustrate these in the context of the recursive model above. As mentioned earlier EViews writes the SVAR system as

$$Az_t = lags + B\eta_t, \tag{4.11}$$

where $A = A_0$, "lags" is $A_1 z_{t-1} + ... + A_p z_{t-p}$. In this form A can be thought of as being used to set up restrictions from behavioral relations (structural equations) and B will be employed for setting up restrictions connected with impulse responses. Because EViews does not allow restrictions on $A_j, j = 1, ..., p$, it is only necessary to specify A and B. That leads us to ignore the lags and to write (4.11)

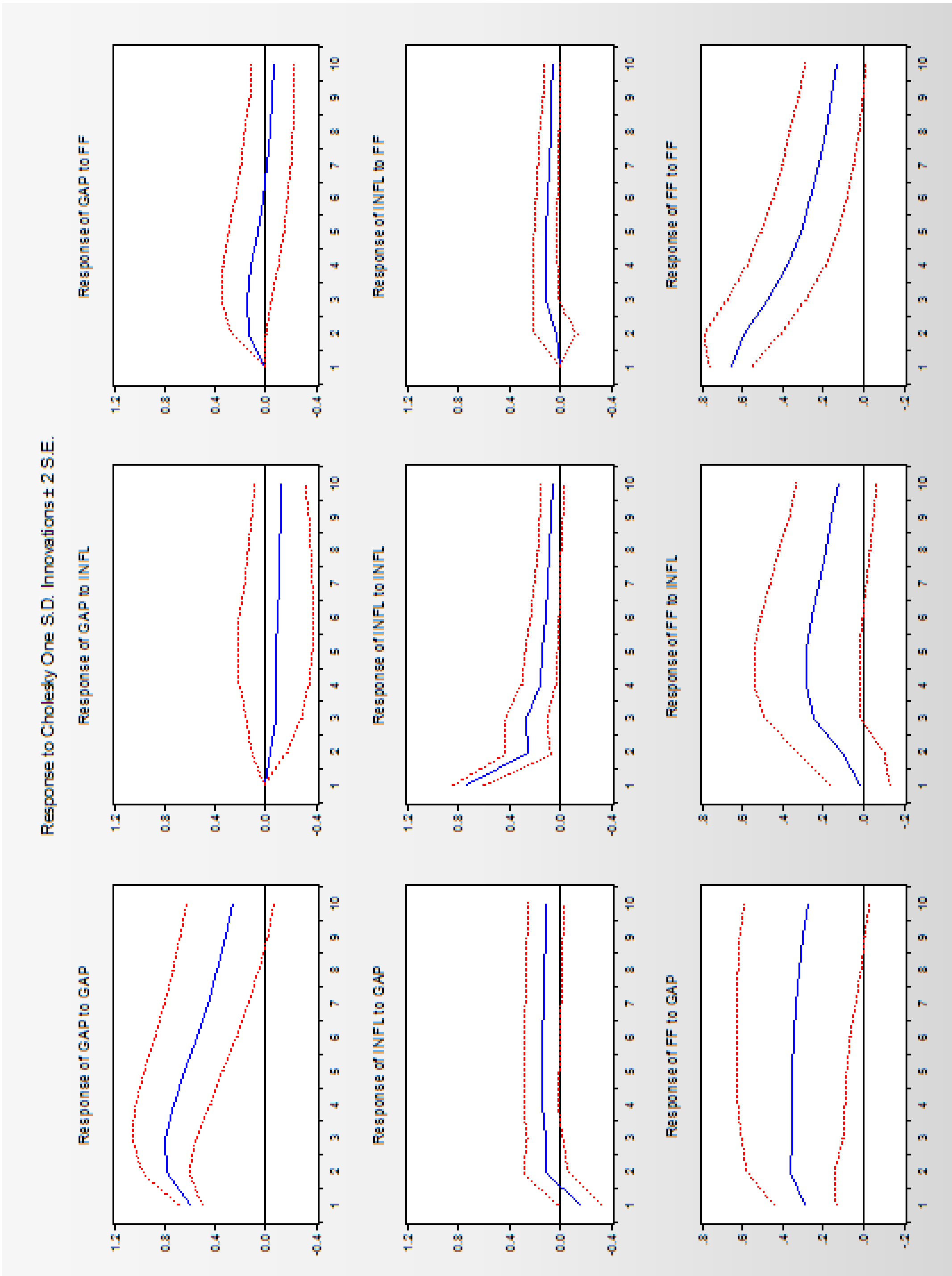


Figure 4.1: Impulse Responses for the Recursive Small Macro Model

as $Ae_t = Bu_t$, where $u_t = \eta_t$ are shocks with unit variance (compared to ε_t which have non-unit variances). The logic of this representation can be seen from an SVAR(1) written in EViews form as $Az_t = A_1 z_{t-1} + B\eta_t$. Substituting the VAR for z_t into this expression the LHS of it can be written as $A(B_1 z_{t-1} + e_t)$. After grouping of terms this becomes $Ae_t = (A_1 - AB_1)z_{t-1} + B\eta_t$. Since $B_1 = A_0^{-1}A_1$ we have $AB_1 = A_0 B_1 = A_1$, leaving $Ae_t = B\eta_t = Bu_t$ in EViews notation. Accordingly, either the matrices A, B or the equations $Ae_t = Bu_t$ need to be provided to EViews, and these two approaches will now be described.

First, return to the screen shot after a VAR(2) has been estimated and select ***Proc→Estimate Structural Factorization***. Then the screen in Figure 4.2 appears and either ***text*** or ***matrix*** must be selected**.** The first of these is used to describe A, B while the second yields $Ae_t = Bu_t$. Dealing with the second of these, a first step is to decide on a normalization. Working with the normalized system in (4.8) - (4.10) and $\varepsilon_{jt} = \sigma_j \eta_{jt} = \sigma_j u_{jt}$, A and B would be

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -a_{21}^0 & 1 & 0 \\ -a_{31}^0 & -a_{32}^0 & 1 \end{bmatrix},\ B = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}.$$

Since we are to write out equations for $Ae_t = Bu_t$, and then get EViews to perform an MLE of the unknown parameters θ in (A_j, B), we need to map the a_{ij}^0 etc. into θ. In EViews the vector θ is described by C and so the mapping might be as follows:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ C(2) & 1 & 0 \\ C(4) & C(5) & 1 \end{bmatrix},\ B = \begin{bmatrix} C(1) & 0 & 0 \\ 0 & C(3) & 0 \\ 0 & 0 & C(6) \end{bmatrix}.$$

Using this characterization the screen shot in Figure 4.2 then shows how $Ae_t = Bu_t$ is written.

Having set up the system to be estimated press **OK.** MLE is then performed and the screen in Figure 4.3 comes up, showing the estimated coefficients in C, the log likelihood and what the estimated A and B matrices look like. To get impulse responses click on ***Impulse→ Impulse Definition→Structural Factorization*.** These are then the same as those found with the Cholesky decomposition.

Now in the second method ***Text*** is chosen rather than ***Matrix*** and this is used to describe A and B directly to EViews. To do so return to the top of the EViews page and create blank matrices using ***Object → New Object → Matrix-Vector Coef***. Select ***OK*** to this, set the number of rows and columns (3 in the small macro model example), and then select **OK.** The screen will then appear as in Figure 4.4.

Edit the spreadsheet on the screen using ***Edit+/-*** so it looks like Figure 4.5 (Here "NA" means that there is an unknown value of the coefficient in the matrix that has to be estimated). Then click on **Name** and call it A.

After doing this repeat the same steps as above to create a B that looks like

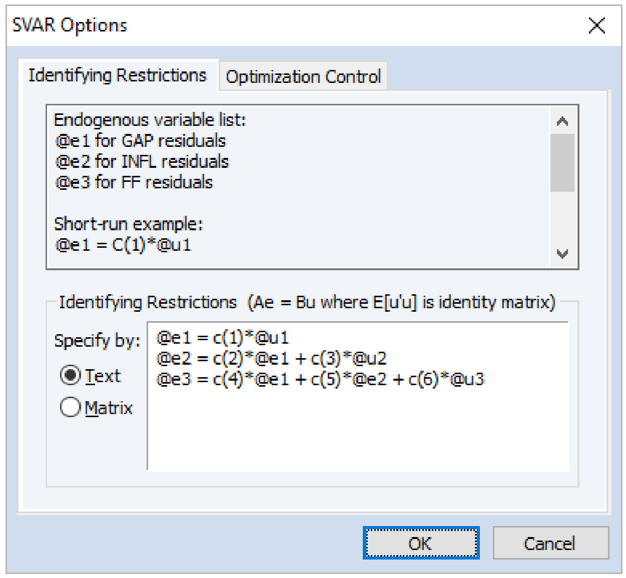


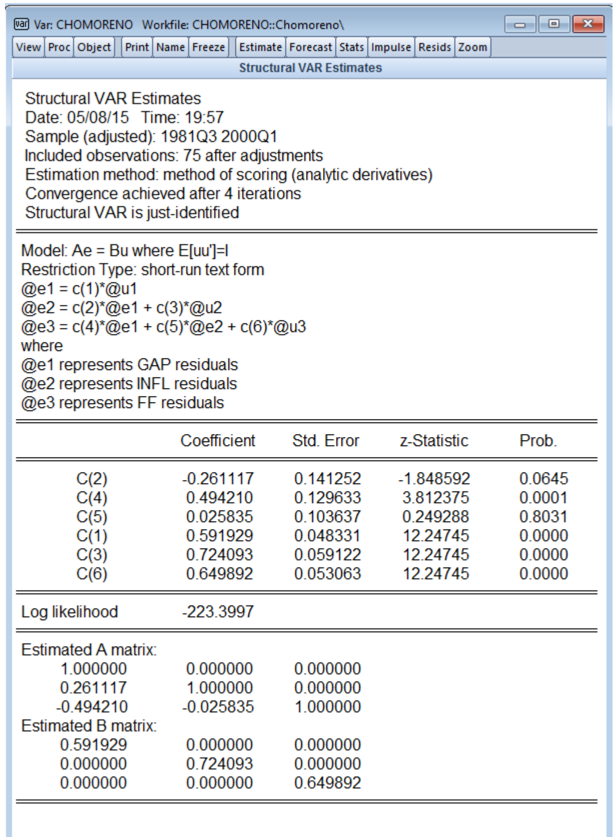
Figure 4.2: Writing $Ae(t) = Bu(t)$ in EViews

$$\begin{array}{cc} & \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{array}{c} R1 \\ R2 \\ R3 \end{array} & \begin{bmatrix} NA & 0 & 0 \\ 0 & NA & 0 \\ 0 & 0 & NA \end{bmatrix} \end{array}$$

As discussed earlier the impulse responses can be used to see which shocks account for the variables at various forecast horizons. To find this information after the impulse responses are computed select ***View*** $\rightarrow$ ***Variance Decomposition*,** filling in the window that is then presented. Using the recursive SVAR(2) fitted to the Cho-Moreno data earlier, the fraction of the variance of inflation ten periods ahead explained by the demand shock is 15.92, by cost shocks is 75.85, and by monetary shocks is 8.24. This is the same result as in Section 4.2 but now the shocks have been given some names.

An alternative approach to estimating the recursive model is estimate Equations 4.8 - 4.10 directly using the System object in EViews. To so, invoke ***Object*** $\rightarrow$ ***New Object***... and complete the resulting dialog box as shown in Figure 4.6. Clicking ***OK*** will create a system object called "chomor_sys" in the workfile. Assuming a base SVAR model with 2 lags, the system object needs to be populated with the EViews code shown in Figure 4.7

The placeholders for the contemporaneous coefficients (a^0_{21}, a^0_{31} and a^0_{32}) are C(22), C(23) and C(24). The next step is to click on the ***Estimate*** tab and the select ***Full Information Maximum Likelihood*** for the estimator using a (restricted) diagonal covariance matrix (see Figure 4.8). The results are shown in Figure 4.9, and match those from the standard SVAR routine



Var: CHOMORENO Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 05/08/15 Time: 19:57
Sample (adjusted): 1981Q3 2000Q1
Included observations: 75 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 4 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1 = c(1)*@u1
@e2 = c(2)*@e1 + c(3)*@u2
@e3 = c(4)*@e1 + c(5)*@e2 + c(6)*@u3
where
@e1 represents GAP residuals
@e2 represents INFL residuals
@e3 represents FF residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(2)	-0.261117	0.141252	-1.848592	0.0645
C(4)	0.494210	0.129633	3.812375	0.0001
C(5)	0.025835	0.103637	0.249288	0.8031
C(1)	0.591929	0.048331	12.24745	0.0000
C(3)	0.724093	0.059122	12.24745	0.0000
C(6)	0.649892	0.053063	12.24745	0.0000

Log likelihood -223.3997

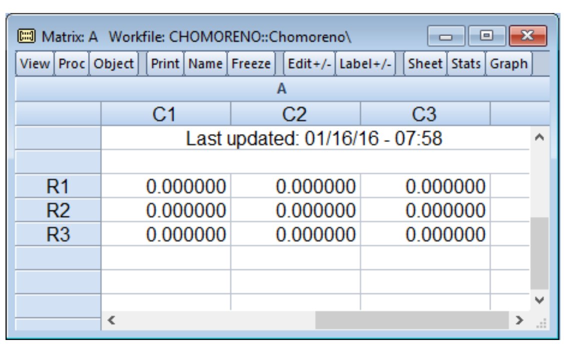
Estimated A matrix:

1.000000	0.000000	0.000000
0.261117	1.000000	0.000000
-0.494210	-0.025835	1.000000

Estimated B matrix:

0.591929	0.000000	0.000000
0.000000	0.724093	0.000000
0.000000	0.000000	0.649892

Figure 4.3: MLE Estimation of A and B matrices for the Small Structural Model Using EViews



Matrix: A Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Edit+/- | Label+/- | Sheet | Stats | Graph

A

	C1	C2	C3
	Last updated: 01/16/16 - 07:58		
R1	0.000000	0.000000	0.000000
R2	0.000000	0.000000	0.000000
R3	0.000000	0.000000	0.000000

Figure 4.4: Example of a Matrix Object in EViews

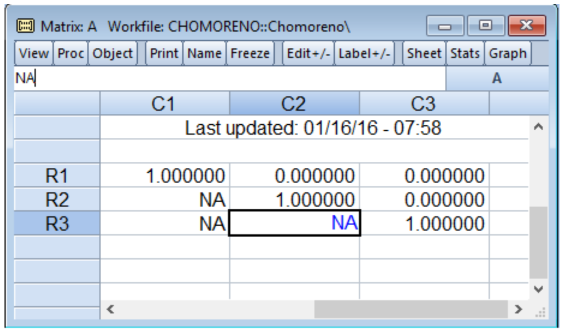


Figure 4.5: Example an A Matrix for the Recursive Small Macro Model

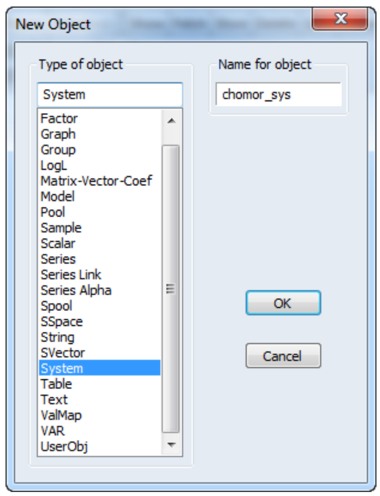


Figure 4.6: Creating a System Object Called *chomor_sys* Using EViews

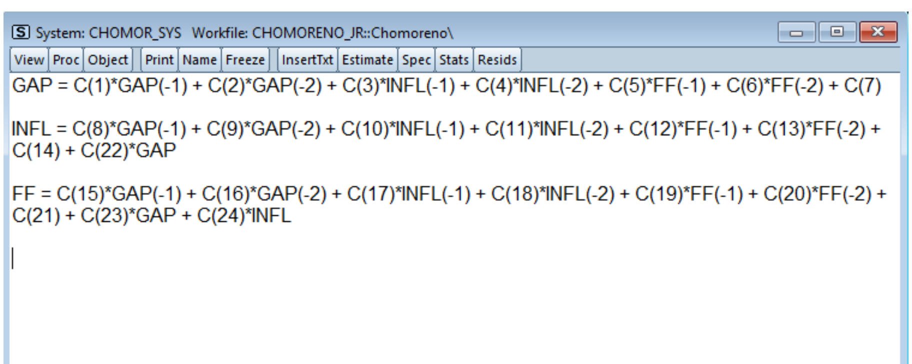


Figure 4.7: Specification of the Small Macro Model in an EViews SYSTEM Object

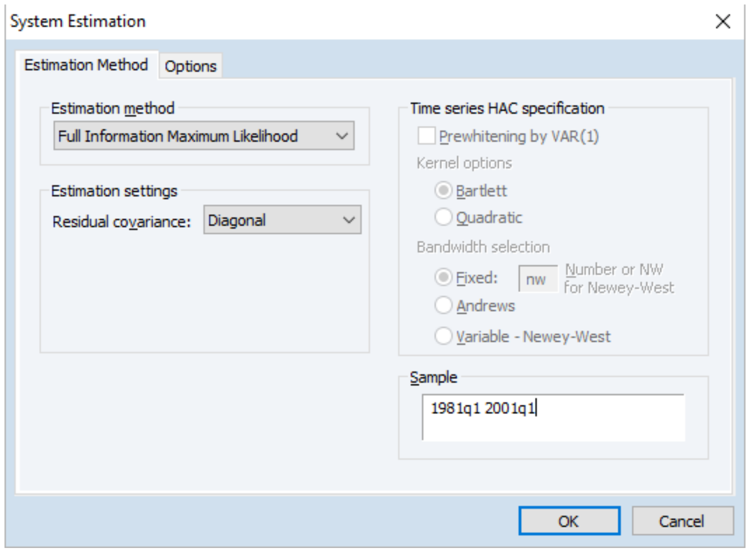


Figure 4.8: System Estimation Using Full Information Maximum Likelihood

in EViews.[4] The resulting A and B matrices are $\begin{bmatrix} 1 & 0 & 0 \\ -C(22) & 1 & 0 \\ -C(23) & -C(24) & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.261117 & 1 & 0 \\ -0.494210 & -0.025836 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0.56363 & 0 & 0 \\ 0 & 0.68947 & 0 \\ 0 & 0 & 0.61882 \end{bmatrix}$ respectively.

Since the model is exactly identified, the implied impulse responses can be computed using the summative VAR as shown in *chomoreno_fiml.prg* (Figure 4.10) using a user specified shock matrix of $A^{-1}B$.[5] The resulting impulse response functions are shown in Figure 4.11 and are identical to those obtained using the standard VAR routine in EViews (see Figure 4.1).

4.6.1.3 Impulse Response Anomalies (Puzzles)

The impulse responses to the interest rate shock in Figure 4.1 show that the responses of inflation and output to an interest rate rise are positive rather than negative as we might have expected. Thus this example is a useful vehicle for making the point that recursive systems often produce “puzzles” such as

1. The *price puzzle* in which monetary policy shocks have a positive effect on inflation.

[4] The standard errors are calculated using the “observed hessian” option of the FIML estimator.

[5] The EViews command is **chomoreno.impulse(10,m,imp=user,se=a,fname=shocks)**, where “shocks” (a matrix in the workfile) $= A^{-1}B$. Doing this shows how it is possible to move from the SYSTEM module back to the SVAR module in order to compute impulse responses.

System: CHOMOR_SYS Workfile: CHOMORENO_JR::Chomoreno\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: CHOMOR_SYS
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Date: 04/12/16 Time: 19:52
Sample: 1981Q3 2000Q1
Included observations: 75
Total system (balanced) observations 225
Residual covariance matrix restricted to be diagonal in FIML estimation
Estimation settings: tol=1.0e-12, derivs=analytic (linear)
Initial Values: C(1)=0.50000, C(2)=0.50000, C(3)=0.50000, C(4)=0.50000, C(5)=0.50000, C(6)=0.50000, C(7)=0.50000, C(8)=0.50000, C(9)=0.50000, C(10)=0.50000, C(11)=0.50000, C(12)=0.50000, C(13)=0.50000, C(14)=0.50000, C(22)=0.50000, C(15)=0.50000, C(16)=0.50000, C(17)=0.50000, C(18)=0.50000, C(19)=0.50000, C(20)=0.50000, C(21)=0.50000, C(23)=0.50000, C(24)=0.50000
Convergence achieved after 72 iterations
Coefficient covariance computed using observed Hessian

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.216215	0.116731	10.41891	0.0000
C(2)	-0.295935	0.113823	-2.599958	0.0093
C(3)	-0.058829	0.091922	-0.639986	0.5222
C(4)	-0.041853	0.088986	-0.470335	0.6381
C(5)	0.193684	0.098304	1.970258	0.0488
C(6)	-0.187740	0.087482	-2.146051	0.0319
C(7)	0.038279	0.067308	0.568716	0.5695
C(8)	0.565720	0.223390	2.532433	0.0113
C(9)	-0.274627	0.145377	-1.889073	0.0589
C(10)	0.331870	0.112752	2.943352	0.0032
C(11)	0.231039	0.109014	2.119347	0.0341
C(12)	0.101246	0.123327	0.820954	0.4117
C(13)	0.010187	0.110253	0.092394	0.9264
C(14)	-0.110674	0.082513	-1.341288	0.1798
C(22)	-0.261117	0.141251	-1.848596	0.0645
C(15)	-0.416907	0.208882	-1.995897	0.0459
C(16)	0.030821	0.133536	0.230810	0.8175
C(17)	0.144704	0.106882	1.353861	0.1758
C(18)	0.191878	0.100729	1.904890	0.0568
C(19)	0.819798	0.111182	7.373484	0.0000
C(20)	-0.060511	0.098957	-0.611486	0.5409
C(21)	-0.166178	0.074940	-2.217468	0.0266
C(23)	0.494210	0.129631	3.812424	0.0001
C(24)	0.025835	0.103632	0.249301	0.8031

Log likelihood	-212.3769	Schwarz criterion	7.044981
Avg. log likelihood	-0.943897	Hannan-Quinn criter.	6.599496

Figure 4.9: FIML Estimates for the Small Macro Model: Diagonal Covariance Matrix


```
'requires chomoreno.wf1
pageselect Chomoreno
smpl @all
'Estimate the summative (reduced form) VAR
var chomoreno.ls 1 2 gap infl ff
chomoreno.results

'Estimate the same model using FIML, a lower-triangular identification
'scheme and a diagonal covariance matrix.

chomor_sys.fiml(covinfo=hessian, rcov=diag)

'Build the contemporaneous A matrix

matrix ahat = @identity(3)
ahat(2,1) = -c(22)
ahat(3,1) = -c(23)
ahat(3,2) = -c(24)

'And the B matrix

matrix bhat = chomor_sys.@estcov
bhat(1,1) = bhat(1,1)^0.5
bhat(2,2) = bhat(2,2)^0.5
bhat(3,3) = bhat(3,3)^0.5

'Since the model is exactly identified, we can use the
'summative model to calculate the impulse response
'functions.

matrix shocks = @inverse(ahat)*bhat
chomoreno.impulse(10,m,imp=user,se=a,fname=shocks)
```

Figure 4.10: EViews Program *chomoreno_fiml.prg* to Calculate Impulse Response Functions

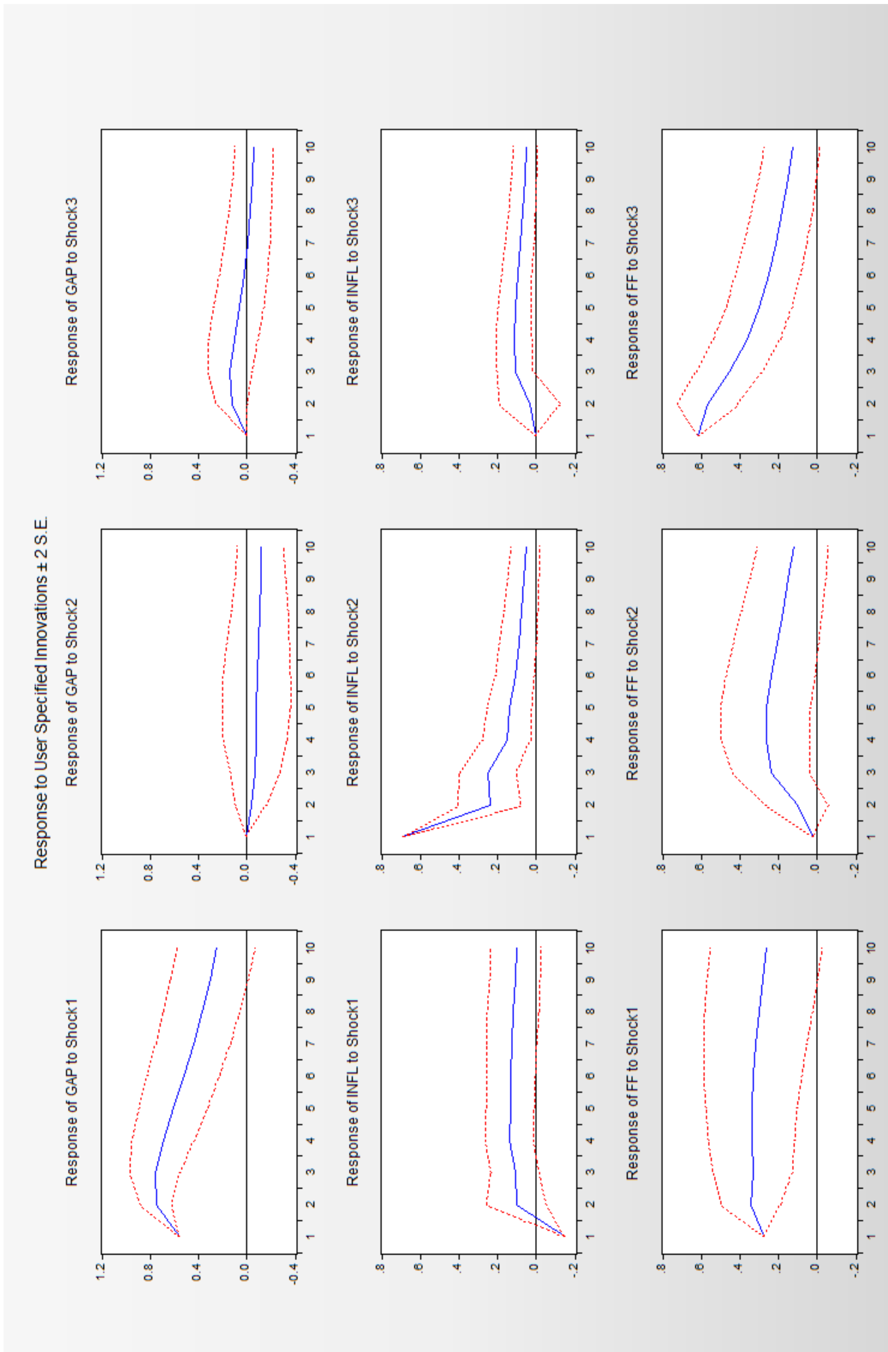


Figure 4.11: Impulse Response Functions for the Recursive Small Macro Model: FIML Estimation with a Diagonal Covariance Matrix

2. The *exchange rate puzzle* in which a monetary shock that raises interest rates depreciates rather than appreciates a currency.

To eliminate such puzzles it is generally necessary to re-specify the SVAR in some way. When variables are stationary there are four approaches to this.

1. Additional variables are added to the system, i.e. there are more shocks. Thus, one variable that is missing from the above system is the stock of money, as there is an implicit money supply equation (interest rate rule) but not a money demand equation. Accordingly, one might add money into the system, which raises the question of what asset demand function would be used, i.e. is it the demand for $M1, M2$, Non-Borrowed Reserves (NBR) or perhaps a Divisia index of money? All of these have been proposed at various times in the literature. There are also other factors that influence policy settings and inflation that might be needed in the system, e.g. oil prices or, more generally, commodity prices. Early studies by Sims (1980) and others did these things, particularly as a way of solving the price puzzle. More recently it has been argued that the addition of factors to the equations can eliminate some puzzles and this will be considered later in the chapter.

2. Re-defining the variables. Giordani (2004) pointed out that it made little sense to use the level of output in an SVAR as the interest rate rule would be expected to depend upon an *output gap rather than the output level*. If the level of output was used, and there is a structural equation linking it to an interest rate, then the growth of output over time would imply larger interest rate movements, unless the coefficient on output in this equation declined. Using the log of output does reduce this effect but does not eliminate it. Moving to an output gap does mean that the coefficient is more likely to be constant and seems closer to what is known about the actual set-up for interest rate decisions, since all theoretical models and institutional studies would suggest that an interest rate rule would involve an output gap and not a level of output. When Giordani used the CBO measure of the U.S. output gap in the SVAR, rather than the level of output, he reduced the price puzzle a great deal. In VARs that have the level of variables such as the log of GDP the addition of a time trend to the exogenous variables means that an approximation to the output gap is being used, where the gap is defined relative to a time trend. But, as we have observed in Chapter 2, adding in the trend to the VAR means that *all* variables will be "detrended" and it is not clear that this is a sensible outcome. In these instances we would prefer to have a time trend in only the structural equation for the log of GDP and not in the other equations. Because the EViews VAR object cannot make a variable exogenous in some equations but not others, programs have to be developed to handle cases where exogenous variables appear in just a sub-set of the structural equations. Another example of this which will be explored later is where there are "external" instruments.

3. Different specifications, e.g. either a non-recursive system or restrictions on the impact of shocks. Kim and Roubini (2000) proposed solving the exchange rate puzzle by allowing a contemporaneous effect of the exchange rate upon the interest rate i.e. the model was no longer recursive.

4. Introducing latent variables so that there are now more shocks than observed variables. The reason for this is that working with a standard SVAR means that the number of shocks equals the number of observed variables, while with latent variables there may be more shocks than observables. If the latent variable is not placed in the system then the impulse responses from the observables-SVAR will be combinations of those for the larger number of shocks, and this may cause difficulties in identifying the shocks of interest. We will not deal specifically with this here but Bache and Leitmo (2008) and Castelnuovo and Surico (2010) present cases where this is the source of the price puzzle. The former have an extra shock to the inflation target while, in the latter, it results from some indeterminacy in the system, i.e. there are "sunspot" shocks.

4.6.2 Imposing Restrictions on the Impact of Shocks

4.6.2.1 A Zero Contemporaneous Restriction

In order to understand some later approaches it is instructive to look at how it is possible to impose the assumption that the contemporaneous impact of a shock upon a variable is zero. For this purpose we will use the small macro model. The relation between the VAR and SVAR (structural) shocks in this three variable case is given by

$$\begin{aligned} e_t &= A_0^{-1}\eta_t = \bar{A}\eta_t \\ &= \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & 0 \\ 0 & \bar{a}_{22} & 0 \\ \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} \end{bmatrix} \eta_t, \end{aligned}$$

where $\bar{A}$ is what is called B in EViews, and the latter will have elements $\{b_{ij}\}$. Now it should be clear that imposing a restriction that $b_{ij} = 0$ in EViews means that e_{it} does not depend on η_{jt}. Consequently contemporaneous restrictions such as

1. Monetary policy shocks (η_{3t}) have a zero contemporaneous effect on output (represented by e_{1t}) and inflation (e_{2t}), and

2. The demand shock (η_{1t}) has a zero contemporaneous effect on inflation,

would mean that

$$e_{1t} = \bar{a}_{11}\eta_{1t} + \bar{a}_{12}\eta_{2t} \tag{4.12}$$

$$e_{2t} = \bar{a}_{22}\eta_{2t} = \varepsilon_{2t} \tag{4.13}$$

$$e_{3t} = \bar{a}_{31}\eta_{1t} + \bar{a}_{32}\eta_{2t} + \bar{a}_{33}\eta_{3t}. \tag{4.14}$$

There are six unknown parameters in this new model and therefore it is exactly identified. It also has exactly the same likelihood as the recursive model fitted earlier to the macro data (and thus is observationally equivalent). Therefore it is not possible to choose between the recursive model and this new one based on fit to the data. Other criteria would be needed to justify selecting one of them.

We illustrate the imposition of the zero restrictions. In terms of the EViews program (4.12) - (4.14) imply that $B = A_0^{-1} = \begin{bmatrix} * & * & 0 \\ 0 & * & 0 \\ * & * & * \end{bmatrix}$, and so the EViews instructions are the same as previously, except that the ***Text*** after ***Estimate*** $\longrightarrow$***Structural Factorization*** is now

```
@e1=c(1)*@u1+c(2)*@u2
@e2=c(3)*@u2
@e3=c(4)*@u1+c(5)*@u2+c(6)*@u3
```

Estimating this model using the SVAR routine yields the output shown in Figure 4.12 in which the implicit A_0 matrix is $\begin{bmatrix} 1 & 0.16689 & 0.0 \\ 0.0 & 1 & 0.0 \\ -0.49429 & -0.02583 & 1 \end{bmatrix}$.[6] Figure 4.13 shows the impulse responses to an interest rate shock from these restrictions. They are very similar to those in Figure 4.1 and, despite the additional restrictions, continue to show price and output puzzles.

Suppose we now think about what the zero impulse response restrictions would imply for a general SVAR. Because $z_{jt} = e_{jt} + lags$ from (4.13) it follows that $z_{2t} = lags + \varepsilon_{2t}$. But ε_{2t} is the structural equation error for z_{2t} so this would imply that this is the structural equation as well, i.e. $a_{21}^0 = 0$, $a_{23}^0 = 0$ meaning that the structural system would look like

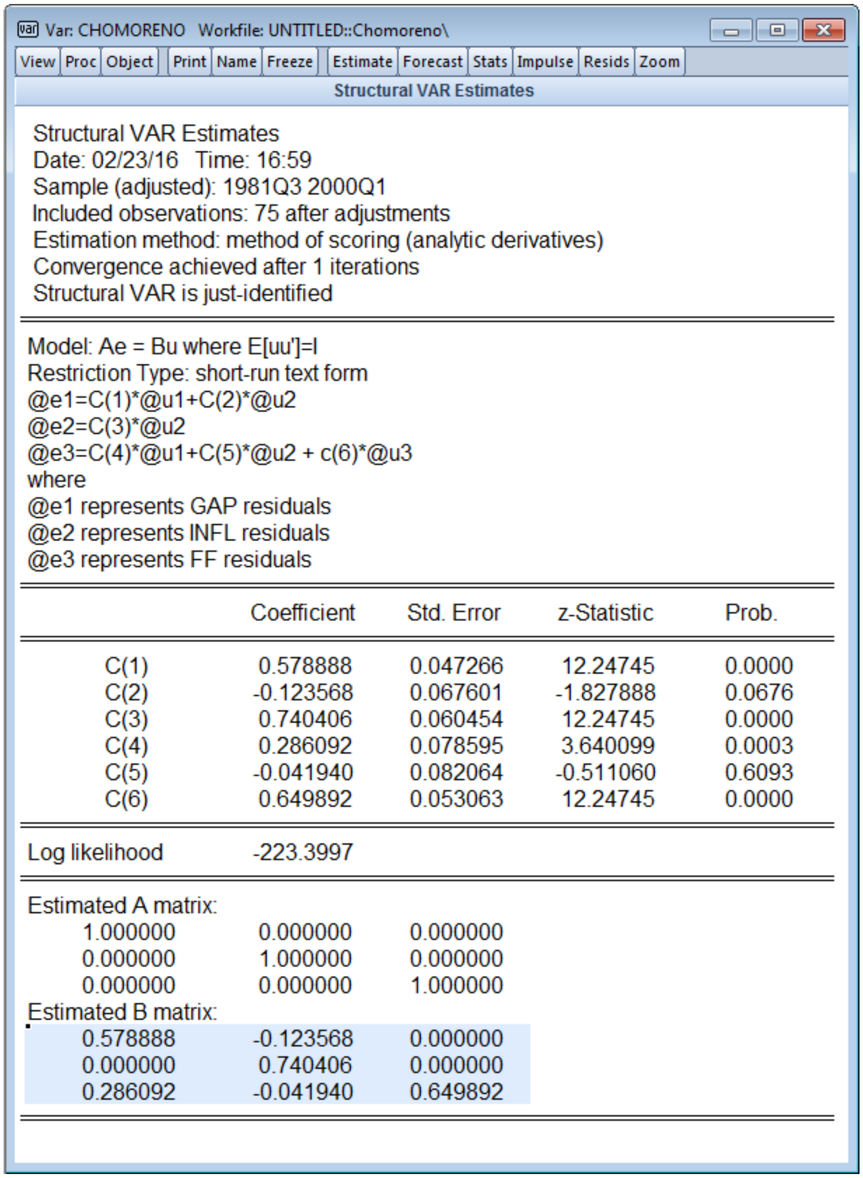
$$z_{1t} = a_{12}^0 z_{2t} + a_{13}^0 z_{3t} + lags + \varepsilon_{1t} \tag{4.15}$$

$$z_{2t} = lags + \varepsilon_{2t} \tag{4.16}$$

$$z_{3t} = a_{31}^0 z_{1t} + a_{32}^0 z_{2t} + lags + \varepsilon_{3t}. \tag{4.17}$$

Accordingly, instruments are needed for the variables in the first and third equations. Starting with the third equation, two instruments are needed that

[6] Note that in this case we are working with an SVAR expressed as $e_t = A_0^{-1}\Gamma u_t$, where Γ is a diagonal matrix containing the standard deviation of the structural errors. The output in Figure 4.18 suggests that the estimated A matrix (i.e., A_0) is the identity matrix. However, EViews is reporting the final estimate of $A_0^{-1}\Gamma$ as the "Estimated B matrix:". The estimate of A_0 can be derived from this quantity as follows. First, the "Estimated B matrix" needs to be inverted, yielding $\mathit{\Gamma}^{-1}A_0$. Second, since the result will not have unity on its diagonal elements it is necessary to divide each row by the diagonal element appearing in it. Thus, after inverting the estimated B matrix in Figure 4.12, to get $\mathit{\Gamma}^{-1}A_0 = \begin{bmatrix} 1.72745 & 0.28830 & 0.0 \\ 0.0 & 1.35061 & 0.0 \\ -0.76045 & -0.03975 & 1.53872 \end{bmatrix}$, renormalization yields $A_0 = \begin{bmatrix} 1 & 0.16689 & 0.0 \\ 0.0 & 1 & 0.0 \\ -0.49429 & -0.02583 & 1 \end{bmatrix}$.


Var: CHOMORENO Workfile: UNTITLED::Chomoreno\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 02/23/16 Time: 16:59
Sample (adjusted): 1981Q3 2000Q1
Included observations: 75 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 1 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1=C(1)*@u1+C(2)*@u2
@e2=C(3)*@u2
@e3=C(4)*@u1+C(5)*@u2 + c(6)*@u3
where
@e1 represents GAP residuals
@e2 represents INFL residuals
@e3 represents FF residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.578888	0.047266	12.24745	0.0000
C(2)	-0.123568	0.067601	-1.827888	0.0676
C(3)	0.740406	0.060454	12.24745	0.0000
C(4)	0.286092	0.078595	3.640099	0.0003
C(5)	-0.041940	0.082064	-0.511060	0.6093
C(6)	0.649892	0.053063	12.24745	0.0000

Log likelihood -223.3997

Estimated A matrix:

1.000000	0.000000	0.000000
0.000000	1.000000	0.000000
0.000000	0.000000	1.000000

Estimated B matrix:

0.578888	-0.123568	0.000000
0.000000	0.740406	0.000000
0.286092	-0.041940	0.649892

Figure 4.12: SVAR Output for the Restricted Small Macro Model

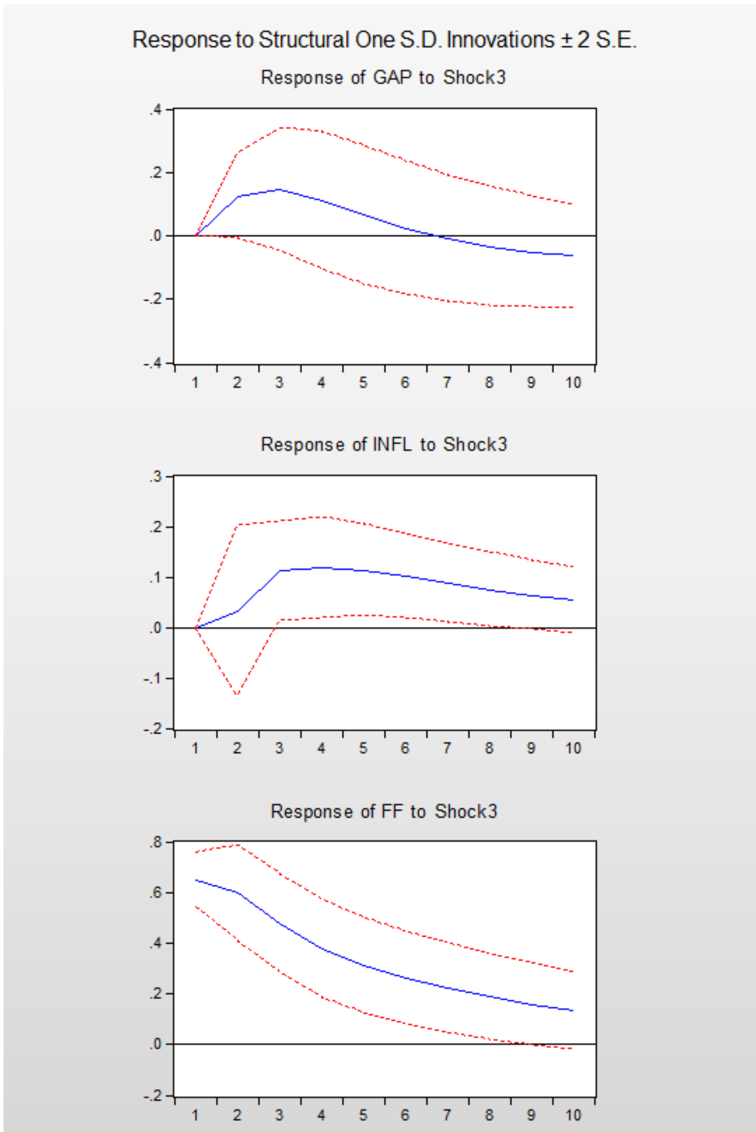


Figure 4.13: Interest Rate Responses from the Small Macro Model Assuming Monetary and Demand Shocks Have Zero Effects

are uncorrelated with ε_{3t}. Looking at (4.12) it is apparent that e_{1t} does not depend on ε_{3t} (as ε_{3t} is a multiple of η_{3t}) and, from (4.13), this is also true of e_{2t}. So e_{1t} and e_{2t} can act as instruments for z_{1t} and z_{2t} in (4.17). Because e_{1t} and e_{2t} are not known it is necessary to use the VAR residuals $\hat{e}_{1t}$ and $\hat{e}_{2t}$ as the instruments. Of course (4.15) can be estimated by OLS since there are no RHS endogenous variables.

In terms of EViews commands it is first necessary to generate the residuals from the VAR equations for z_{1t} and z_{2t}. While these residuals can be obtained from the VAR output they can also be found by the commands ***Quick*** $\rightarrow$ ***Estimate Equation*** and then choosing ***LS - Least Squares (NLS and ARMA)***. The specification box needs to be filled in with

$$gap\ gap(-1)\ gap(-2)\ infl(-1)\ infl(-2)\ ff(-1)\ ff(-2)$$

Clicking on ***OK*** then gives the parameter estimates. To construct the residuals $\hat{e}_{1t}$ from this regression and save them in the workfile, click on ***Proc***$\rightarrow$***Make Residual Series*** and then ***OK*** after giving the residuals a name of "res1". Repeating this process but with the dependent variable being *infl* will give the residuals $\hat{e}_{2t} = \hat{\varepsilon}_{2t}$ (the VAR and structural errors are the same as there are no RHS endogenous variables in this equation). Note that these residuals are automatically saved in a series called "*eps2*" by the program shown in Figure 4.14.

As detailed above, $\hat{e}_{1t}$ and $\hat{\varepsilon}_{2t}$ will be the instruments for z_{1t} and z_{2t} in the third (interest rate) equation. This can be done from the screen presented at the end of the OLS estimates. Choose ***Estimate*** from the options available and then fill in the specification as

$$ff\ gap\ infl\ gap(-1)\ gap(-2)\ infl(-1)\ infl(-2)\ ff(-1)\ ff(-2)$$

Instead of selecting the LS option, choose ***TSLS - Two-Stage Least Squares (TSNLS and ARMA)***, whereupon it will ask for ***Instrument List*** to be filled in. Insert

$$res1\ eps2\ gap(-1)\ gap(-2)\ infl(-1)\ infl(-1)\ ff(-1)\ ff(-1).$$

After the IV estimates are obtained the command ***Proc***$\rightarrow$***Make Residual Series*** can be used to create a series object containing the residuals from this equation. Suppose this series is called "*eps3*". The procedure is then repeated for the first equation, with the model defined as

$$gap\ infl\ ff\ gap(-1)\ gap(-2)\ infl(-1)\ infl(-2)\ ff(-1)\ ff(-2),$$

and the instruments being

$$eps2\ eps3\ gap(-1)\ gap(-2)\ infl(-1)\ infl(-2)\ ff(-1)\ ff(-2).$$

The resulting parameter estimates are $A_0 = \begin{bmatrix} 1 & .1669 & 0 \\ 0 & 1 & 0 \\ -.494 & -.0258 & 1 \end{bmatrix}$ and

$B = \begin{bmatrix} .5788 & 0 & 0 \\ 0 & .7404 & 0 \\ 0 & 0 & .6596 \end{bmatrix}$, where the diagonal elements of B are the estimated standard deviations of the errors of the equations. The IV parameter estimates are identical to those reported in Figure 4.12 from the SVAR routine. Lastly, instead of using the pull-down menus above to get the instrumental variable results we can build an EViews program that will do this. The code, which is saved in *chomoreno_restrict.prg,* is shown in Figure 4.14.

An important point to remember from this application is that if the $l'th$ shock has a zero contemporaneous effect on the $k'th$ variable it means that the $k'th$ equation VAR residuals can be used as instruments in estimating the $l'th$ structural equation ("the VAR instrument principle"). We will utilize this result many times in the material that follows.

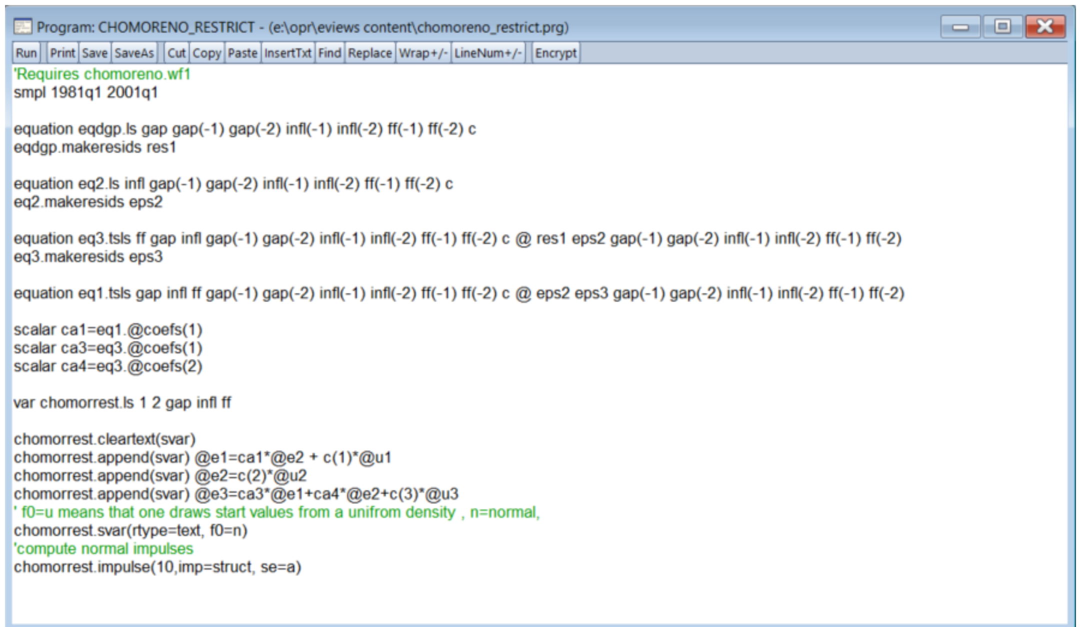
An interesting feature of the IV approach in this case is that it automatically imposes an implicit constraint of $a^0_{13} = 0$ on the structural VAR. This constraint ensures that e_{1t} (the VAR residual of the first equation) is not affected by η_{3t}. By construction, both e_{2t} and e_{3t} (the VAR residuals) are orthogonal to z_{3t}. After incorporating the lagged variables as instruments, z_{3t} will therefore be instrumented by $z_{3t} - lags = \varepsilon_{3t}$, which by assumption does not affect z_{1t}. Another way of seeing that $a^0_{13} = 0$ is to invert $A_0^{-1} = \begin{bmatrix} * & * & 0 \\ 0 & * & 0 \\ * & * & * \end{bmatrix}$ symbolically. Doing so reveals that A_0 will have the same structure as A_0^{-1} in terms of the position of the zero elements. It is also clear that the equations could be re-arranged into a recursive structure. Hence $a^0_{13} = 0$.

The VAR instrument principle may also be used in the SYSTEM estimator by re-specifying the model so that it incorporates the VAR residuals explicitly along with the binding constraint that $a^0_{13} = 0$, viz:

$$\begin{aligned} z_{1t} &= a^0_{12}(z_{2t} - lags) + lags + \varepsilon_{1t} && (4.18) \\ z_{2t} &= lags + \varepsilon_{2t} && (4.19) \\ z_{3t} &= a^0_{31}(z_{1t} - lags) + a^0_{32}(z_{2t} - lags) + lags + \varepsilon_{3t}. && (4.20) \end{aligned}$$

The necessary EViews code is given in Figure 4.15, with C(22), C(23) and C(24) corresponding to the contemporaneous parameter estimates a^0_{12}, a^0_{31} and a^0_{32}. Estimating the system object (*ch_sys_iv_rest*) using ordinary least squares yields the output shown in Figure 4.16. The estimates for a^0_{12}, a^0_{31} and a^0_{32} match those obtained using the instrumental variable approach.

Lastly, one may also estimate the restricted system using FIML and the diagonal covariance matrix option. The required system object code (see *ch_sys_iv_rest* in the workfile) is shown in Figure 4.17 and the results, which match the IV estimates, are shown in Figure 4.18




```
'Requires chomoreno.wf1
smpl 1981q1 2001q1

equation eqdgp.ls gap gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2) c
eqdgp.makeresids res1

equation eq2.ls infl gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2) c
eq2.makeresids eps2

equation eq3.tsls ff gap infl gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2) c @ res1 eps2 gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2)
eq3.makeresids eps3

equation eq1.tsls gap infl ff gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2) c @ eps2 eps3 gap(-1) gap(-2) infl(-1) infl(-2) ff(-1) ff(-2)

scalar ca1=eq1.@coefs(1)
scalar ca3=eq3.@coefs(1)
scalar ca4=eq3.@coefs(2)

var chomorrest.ls 1 2 gap infl ff

chomorrest.cleartext(svar)
chomorrest.append(svar) @e1=ca1*@e2 + c(1)*@u1
chomorrest.append(svar) @e2=c(2)*@u2
chomorrest.append(svar) @e3=ca3*@e1+ca4*@e2+c(3)*@u3
' f0=u means that one draws start values from a unifrom density , n=normal,
chomorrest.svar(rtype=text, f0=n)
'compute normal impulses
chomorrest.impulse(10,imp=struct, se=a)
```

Figure 4.14: *chomoreno_restrict.prg* to Perform IV on the Restricted Model and Calculate Impulse Responses

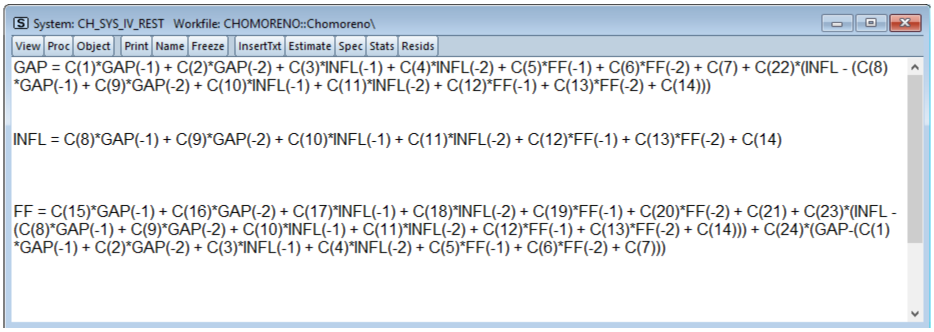


Figure 4.15: EViews System Specification For Equations 4.18 - 4.20

4.6.2.2 A Two Periods Ahead Zero Restriction

We want to impose a restriction on the impulse responses of a variable to a shock at some lag length other than zero. The method is that used in McKibbin *et al.* (1998). It uses the result from (4.5) that $C_j = D_j C_0$. Because D_j are just impulse responses from the VAR they can be found without reference to a structure. So, if the SVAR is set up as $Az_t = lags + B\eta_t$, where η_t have unit variances (the η_t are then EViews' u_t), putting $A = I$ shows that $B = C_0$. B will of course contain one standard deviation shocks. Hence in terms of the A, B structure $C_j = D_j B$.

Now suppose it is desired to impose that the first variable response to the third shock is zero at the second horizon. This will mean that

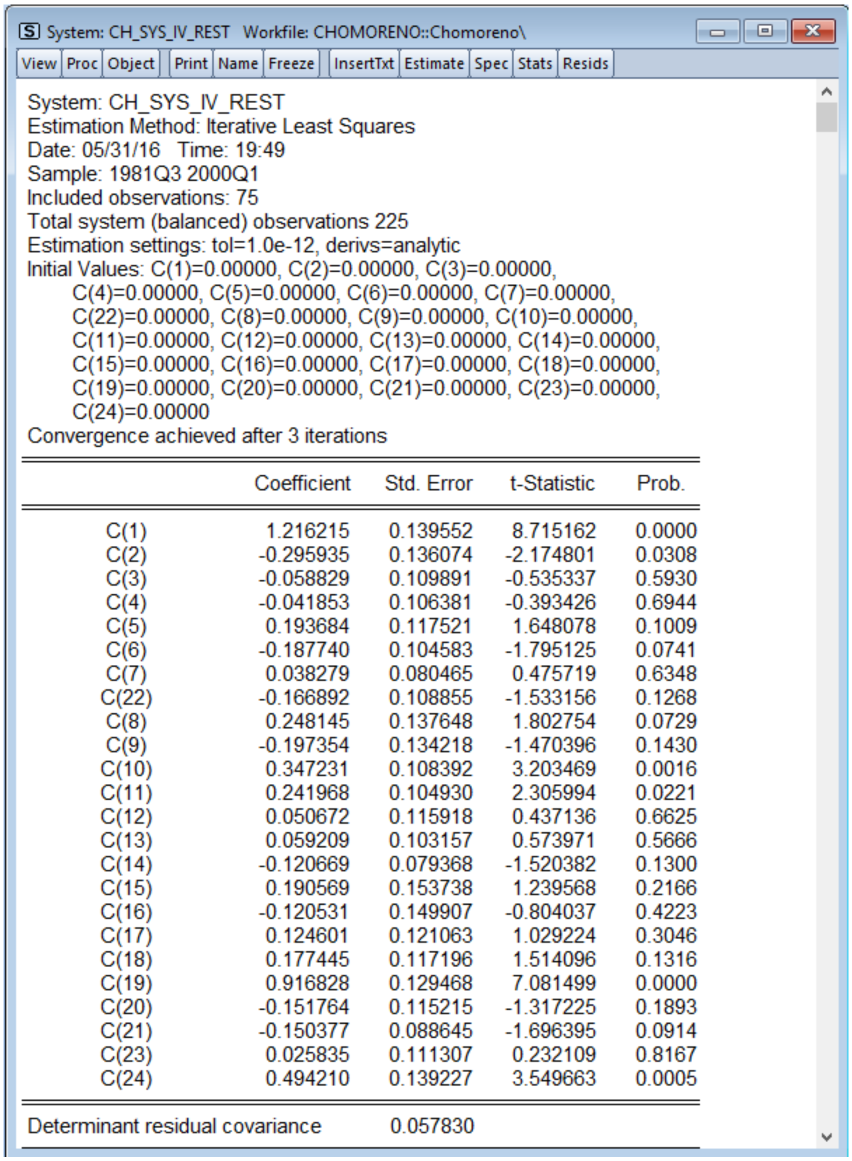
$$[C_2]_{13} = [D_2 B]_{13} = 0,$$

where $[F]_{ij}$ refers to the i, j'th element of the matrix F. Consequently

$$c^2_{13} = \begin{bmatrix} d^2_{11} & d^2_{12} & d^2_{13} \end{bmatrix} \begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = 0$$

is the restriction to be imposed. Clearly this implies $b_{13}d^2_{11} + b_{23}d^2_{12} + b_{33}d^2_{13} = 0$. Because d^2_{ij} are known from the estimated VAR this provides a linear restriction on the elements of B, which is easy to apply in EViews by using the text form of restrictions.

We use the Cho and Moreno data set, fit a VAR(2) with the variables gap, $infl$, ff, and recover D (as was done in Chapter 3). Figure 3.20 gave the responses of the output gap to the VAR residuals for two periods ahead and these provide d^2_{ij}. These were $d^2_{11} = 1.205556; d^2_{12} = -.109696; d^2_{13} = .222415$ giving the restriction $1.205556b_{13} - .109696b_{23} + .222415b_{33} = 0$. This can then be used to substitute out for b_{33}. Now three restrictions on B are needed to estimate any SVAR with three variables. This gives us one. So we need two more. Because we don't want b_{13}, b_{23} and b_{33} all to be zero (we could put one to



System: CH_SYS_IV_REST Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: CH_SYS_IV_REST
Estimation Method: Iterative Least Squares
Date: 05/31/16 Time: 19:49
Sample: 1981Q3 2000Q1
Included observations: 75
Total system (balanced) observations 225
Estimation settings: tol=1.0e-12, derivs=analytic
Initial Values: C(1)=0.00000, C(2)=0.00000, C(3)=0.00000,
C(4)=0.00000, C(5)=0.00000, C(6)=0.00000, C(7)=0.00000,
C(22)=0.00000, C(8)=0.00000, C(9)=0.00000, C(10)=0.00000,
C(11)=0.00000, C(12)=0.00000, C(13)=0.00000, C(14)=0.00000,
C(15)=0.00000, C(16)=0.00000, C(17)=0.00000, C(18)=0.00000,
C(19)=0.00000, C(20)=0.00000, C(21)=0.00000, C(23)=0.00000,
C(24)=0.00000
Convergence achieved after 3 iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.216215	0.139552	8.715162	0.0000
C(2)	-0.295935	0.136074	-2.174801	0.0308
C(3)	-0.058829	0.109891	-0.535337	0.5930
C(4)	-0.041853	0.106381	-0.393426	0.6944
C(5)	0.193684	0.117521	1.648078	0.1009
C(6)	-0.187740	0.104583	-1.795125	0.0741
C(7)	0.038279	0.080465	0.475719	0.6348
C(22)	-0.166892	0.108855	-1.533156	0.1268
C(8)	0.248145	0.137648	1.802754	0.0729
C(9)	-0.197354	0.134218	-1.470396	0.1430
C(10)	0.347231	0.108392	3.203469	0.0016
C(11)	0.241968	0.104930	2.305994	0.0221
C(12)	0.050672	0.115918	0.437136	0.6625
C(13)	0.059209	0.103157	0.573971	0.5666
C(14)	-0.120669	0.079368	-1.520382	0.1300
C(15)	0.190569	0.153738	1.239568	0.2166
C(16)	-0.120531	0.149907	-0.804037	0.4223
C(17)	0.124601	0.121063	1.029224	0.3046
C(18)	0.177445	0.117196	1.514096	0.1316
C(19)	0.916828	0.129468	7.081499	0.0000
C(20)	-0.151764	0.115215	-1.317225	0.1893
C(21)	-0.150377	0.088645	-1.696395	0.0914
C(23)	0.025835	0.111307	0.232109	0.8167
C(24)	0.494210	0.139227	3.549663	0.0005

Determinant residual covariance 0.057830

Figure 4.16: Non-linear Least Squares Estimates of Equations 4.18 - 4.20

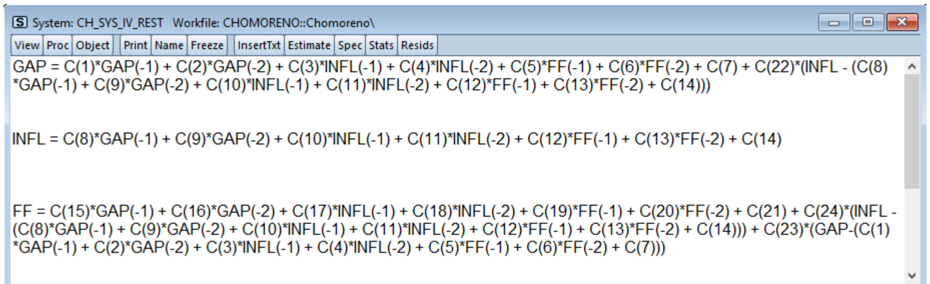


Figure 4.17: EViews System Specification For Equations 4.15 - 4.17 Assuming $a_{13} = 0$.

zero if we wanted) it makes sense to set $b_{12} = 0, b_{21} = 0$. Then the text form of the code to estimate the structure corresponding to the model that incorporates the second period restriction will be

```
@e1=C(1)*@u1+C(2)*@u3
@e2=C(3)*@u2+C(4)*@u3
@e3=C(5)*@u1+C(6)*@u2-(1.0/0.222415)*(1.205556*C(2)-0.109696*C(4))*@u3
```

and the estimated impulse responses in Figure 4.19 confirm that this approach imposes the required restriction (see Shock3, period 3[7]).

We can use the same approach to impose restrictions on the cumulative sum of impulses, $\sum_{j=1}^{P} C_j$, since $\sum_{j=1}^{P} C_j = (\sum_{j=1}^{P} D_j)B$. If $P = 9$ and $\sum_{j=1}^{P} c_{13}^{j} = 0$, then the accumulated values of D_j up to the $P'th$ lag can be found from the VAR impulse responses. For the example that has just been done these will be

$$\sum_{j=1}^{P} d_{11}^{j} = 8.756985, \sum_{j=1}^{P} d_{12}^{j} = -1.176505, \sum_{j=1}^{P} d_{13}^{j} = .480120,$$

and then the third text command would now be

```
@e3=C(5)*@u1+C(6)*@u2-(1.0/0.480120)*(8.756985*C(2)-1.176505*C(4))*@u3
```

forcing the accumulated impulse $\sum_{j=1}^{P} c_{13}^{j}$ to be zero.

4.6.3 Imposing Restrictions on Parameters - The Blanchard-Perotti Fiscal Policy Model

Blanchard and Perotti (2002) are interested in finding out what the impact of spending and taxes are on GDP. They have three variables z_{1t} = log of real per capita taxes, z_{2t} = log of real per capita expenditures and z_{3t} = log of real per

[7] "Period 3" corresponds to the response two periods ahead as EViews refers to the contemporaneous (zero-period ahead) as "1".

System: CH_SYS_IV_REST Workfile: CHOMORENO::Chomoreno\

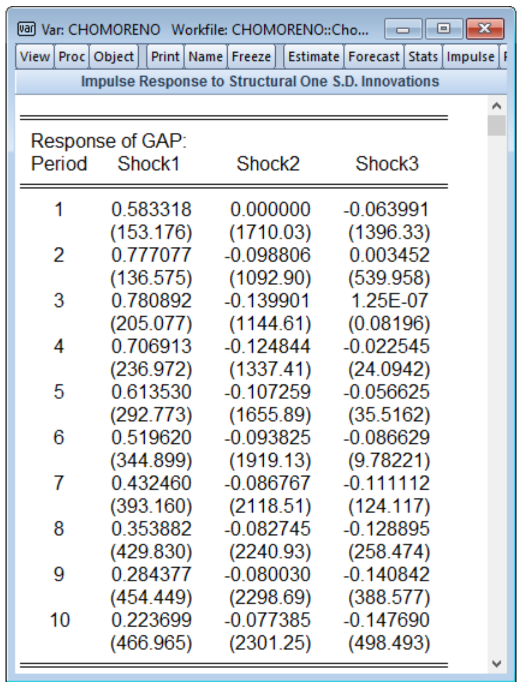
View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: CH_SYS_IV_REST
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Date: 05/31/16 Time: 19:41
Sample: 1981Q3 2000Q1
Included observations: 75
Total system (balanced) observations 225
Residual covariance matrix restricted to be diagonal in FIML estimation
Estimation settings: tol=1.0e-16, derivs=analytic
Initial Values: C(1)=0.00000, C(2)=0.00000, C(3)=0.00000, C(4)=0.00000, C(5)=0.00000, C(6)=0.00000, C(7)=0.00000, C(22)=0.00000, C(8)=0.00000, C(9)=0.00000, C(10)=0.00000, C(11)=0.00000, C(12)=0.00000, C(13)=0.00000, C(14)=0.00000, C(15)=0.00000, C(16)=0.00000, C(17)=0.00000, C(18)=0.00000, C(19)=0.00000, C(20)=0.00000, C(21)=0.00000, C(23)=0.00000, C(24)=0.00000
Convergence achieved after 51 iterations
Coefficient covariance computed using observed Hessian

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.216215	0.116731	10.41893	0.0000
C(2)	-0.295935	0.113823	-2.599962	0.0093
C(3)	-0.058829	0.091922	-0.639986	0.5222
C(4)	-0.041853	0.088985	-0.470339	0.6381
C(5)	0.193684	0.098303	1.970266	0.0488
C(6)	-0.187740	0.087481	-2.146057	0.0319
C(7)	0.038279	0.067307	0.568718	0.5695
C(22)	-0.166892	0.090281	-1.848593	0.0645
C(8)	0.248145	0.146012	1.699490	0.0892
C(9)	-0.197354	0.142374	-1.386168	0.1657
C(10)	0.347231	0.114979	3.019956	0.0025
C(11)	0.241968	0.111306	2.173900	0.0297
C(12)	0.050672	0.122959	0.412103	0.6803
C(13)	0.059209	0.109423	0.541100	0.5884
C(14)	-0.120669	0.084190	-1.433289	0.1518
C(15)	0.190569	0.140275	1.358538	0.1743
C(16)	-0.120531	0.136780	-0.881206	0.3782
C(17)	0.124601	0.110461	1.128003	0.2593
C(18)	0.177445	0.106933	1.659412	0.0970
C(19)	0.916828	0.118130	7.761203	0.0000
C(20)	-0.151764	0.105125	-1.443654	0.1488
C(21)	-0.150377	0.080883	-1.859208	0.0630
C(23)	0.025836	0.103638	0.249287	0.8031
C(24)	0.494210	0.129633	3.812375	0.0001

Log likelihood	-212.3769	Schwarz criterion	7.044981
Avg. log likelihood	-0.943897	Hannan-Quinn criter.	6.599496
Akaike info criterion	6.303385		
Determinant residual covariance		0.057830	

Figure 4.18: FIML Estimates of Equations 4.15 - 4.17 Assuming $a_{13} = 0$.



Response of GAP:

Period	Shock1	Shock2	Shock3
1	0.583318	0.000000	-0.063991
	(153.176)	(1710.03)	(1396.33)
2	0.777077	-0.098806	0.003452
	(136.575)	(1092.90)	(539.958)
3	0.780892	-0.139901	1.25E-07
	(205.077)	(1144.61)	(0.08196)
4	0.706913	-0.124844	-0.022545
	(236.972)	(1337.41)	(24.0942)
5	0.613530	-0.107259	-0.056625
	(292.773)	(1655.89)	(35.5162)
6	0.519620	-0.093825	-0.086629
	(344.899)	(1919.13)	(9.78221)
7	0.432460	-0.086767	-0.111112
	(393.160)	(2118.51)	(124.117)
8	0.353882	-0.082745	-0.128895
	(429.830)	(2240.93)	(258.474)
9	0.284377	-0.080030	-0.140842
	(454.449)	(2298.69)	(388.577)
10	0.223699	-0.077385	-0.147690
	(466.965)	(2301.25)	(498.493)

Figure 4.19: Impulse Responses of the Output Gap to Supply, Demand and Monetary Shocks

capita GDP. The SVAR model has the form (see Blanchard and Perotti (2004, p 1333))

$$\begin{aligned} z_{1t} &= a_1 z_{3t} + a_2' \varepsilon_{2t} + lags + \varepsilon_{1t} \\ z_{2t} &= b_1 z_{3t} + b_2' \varepsilon_{1t} + lags + \varepsilon_{2t} \\ z_{3t} &= \delta_1 z_{1t} + \delta_2 z_{2t} + lags + \varepsilon_{3t} \end{aligned} \tag{4.21}$$

Accordingly, in the EViews representation the A and B matrices have the form

$$A = \begin{bmatrix} 1 & 0 & -a_1 \\ 0 & 1 & -b_1 \\ -\delta_1 & -\delta_2 & 1 \end{bmatrix}, \; B = \begin{bmatrix} * & a_2' & 0 \\ b_2' & * & 0 \\ 0 & 0 & * \end{bmatrix}.$$

A value of 2.08 is given to a_1 by noting that it is the elasticity of taxes with respect to (w.r.t.) GDP. That quantity can be decomposed as the product of the elasticity w.r.t. the tax base and the elasticity of the tax base w.r.t. GDP. These elasticities are computed for a range of taxes and then aggregated to produce a value for a_1. The parameter b_1 is set to zero since they say "We could not identify any automatic feedback from economic activity to government purchases..." (p 1334). Because $n = 3$ only six parameters in A and B can be estimated. However, after fixing a_1 to 2.08, and b_1 to zero, seven unknown parameters remain. This means that one of a_2' or b_2' needs to be prescribed. Looking at the coefficients in B we see that the (2,1) element (b_2') is the response of expenditure to a structural shock in taxes within the quarter, while the (1,2) element (a_2') is how taxes respond to expenditures. Blanchard and Perotti sequentially set either a_2' or b_2' to zero, estimating the other one. We will perform estimation with $b_2' = 0$.

In terms of structural equations consider what we have (when $b_2^{'} = 0$)[8]

$$z_{1t} = 2.08 z_{3t} + a_2' \eta_{2t} + lags + \sigma_1 \eta_{1t} \tag{4.22}$$

$$z_{2t} = lags + \sigma_2 \eta_{2t} = lags + \varepsilon_{2t} \tag{4.23}$$

$$z_{3t} = \delta_1 z_{1t} + \delta_2 z_{2t} + lags + \sigma_3 \eta_{3t} \tag{4.24}$$

Now, because v_t= $\mathrm{z}_{it} - 2.08 z_{3t}$ does not depend on η_{3t}, Blanchard and Perotti used it as an instrument for z_{1t} in (4.24) while z_{2t} could be used as an instrument for itself. This gave them estimates of δ_1 and $\delta_2$and produced the shock ε_{3t}. To estimate the remaining parameters of the system one could estimate $a_2^{'}$ by regressing v_t on lagged values and $\hat{\varepsilon_{2t}} = \hat{\sigma}_2 \hat{\eta}_{2t}$. This is what we will refer to as their IV strategy.

A complication now arises in their work since they added extra regressors to the equations as control variables. These involved dummies for the temporary tax rebate of 1975:2, a quadratic polynomial in time, and an allowance for seasonal variation in the lag coefficients. To do the latter it is necessary to

[8] $a_2 = a_2' \sigma_1$.

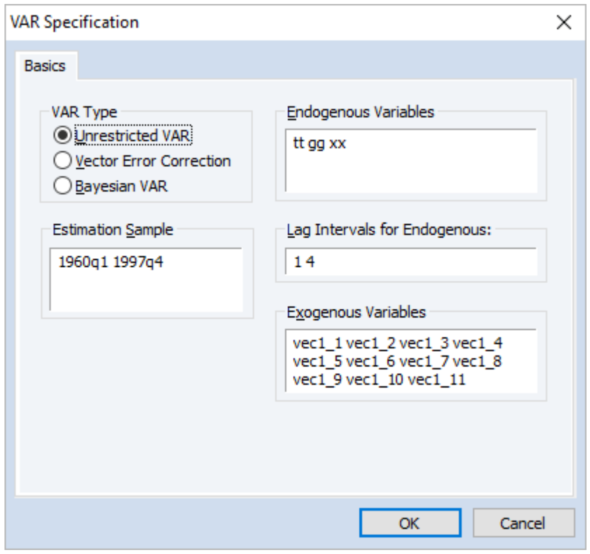


Figure 4.20: Instrumental Variable Estimation of the Blanchard-Perotti Model using EViews

construct multiplicative variables such as $z_{1t-j} \times S_{kt}$ $(k = 1, .., 4)$. Blanchard and Perotti add all these variables to the third equation to get the IV estimates but, when they come to estimating the remaining coefficients in the SVAR, they only add in the tax cut dummy, seasonal intercept shifts, and the quadratic polynomials, thereby preventing the the impulse responses from varying with the seasons. Hence the regressors "vec_*" below represent the last mentioned variables (and because the extended set of variables effectively incorporate an intercept, the constant needs to be removed from the equations as well as from the instrument set). Because of the treatment of these extended regressors there will be a difference between Blanchard and Perotti's IV and the MLE estimates of δ_1 and δ_2. To find the latter first fix a_1 and $b_2^{'}$ yielding

$$A = \begin{bmatrix} 1 & 0 & -2.08 \\ 0 & 1 & 0 \\ * & * & 1 \end{bmatrix}, \; B = \begin{bmatrix} * & * & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix}.$$

Using the workfile *bp.wf1* the screens in Figures 4.20 and 4.21 show how this model is estimated. The resulting MLE estimates of A and B are:[9]

[9] The IV estimates of δ_1 and δ_2 would be -0.134 and 0.236 and, as expected, will differ from the MLE. It should be noted that one cannot add on all the regressors as exogenous variables in the SVAR with a pull-down menu as there is an upper limit to the number.

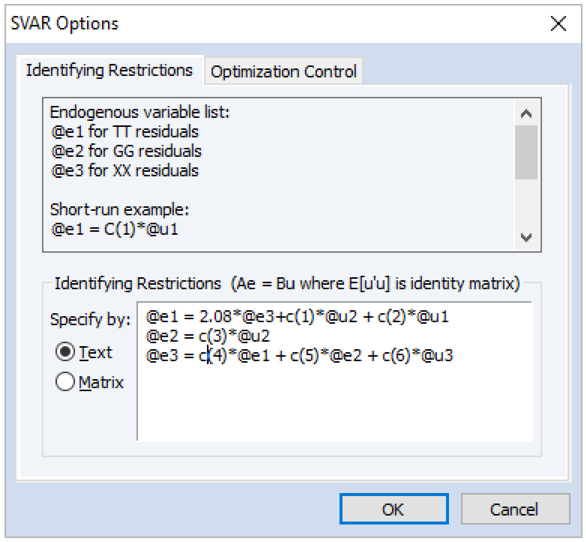


Figure 4.21: Structural VAR Estimation of the Blanchard-Perotti Model using EViews

$$A = \begin{bmatrix} 1 & 0 & -2.08 \\ 0 & 1 & 0 \\ .1343 & -.2879 & 1 \end{bmatrix}, B = \begin{bmatrix} .2499 & -.0022 & 0 \\ 0 & .0098 & 0 \\ 0 & 0 & .0096 \end{bmatrix}.$$

4.6.4 Incorporating Stocks and Flows Plus Identities into an SVAR - A US Fiscal-Debt Model

As well as re-specification in an attempt to eliminate puzzles it also needs to be recognized that SVARs may need to introduce extra items that are routinely present in macro models. Foremost among these are identities. Thus, if an inflation rate $\pi_t = \Delta log(P_t)$ appears in an SVAR, to get the impact on the price level it is necessary to use the identity $log(P_t) = log(P_{t-1}) + \pi_t$. If only π_t (and not $log(P_t)$) enters the SVAR then the impulse responses for $log(P_t)$ can be found by accumulating those for inflation. However there is recent work that argues for the deviation of the price level from its target path to appear in the interest rate rule, i.e. a term like $(log(P_t) - \bar{\pi}t)$ should be present, as well as inflation and an output gap. In such an SVAR both π_t and $log(P_t)$ are present among the variables, and the identity linking $log(P_t)$ and π_t will need to be imposed. For a number of reasons this can't be done in a standard way using the SVAR routine in EViews, since it assumes that there are the same number of shocks as observed variables. However, when there is an identity in the system, the shock for that equation is zero. One can substitute out a variable that is defined by a static identity. As we saw earlier the system remains a VAR in the smaller number of variables. If however the identity is dynamic then the situation is more complex, and we provide a workaround in what follows.

Dynamic identities come up fairly frequently: an example would be when stock variables such as household assets are introduced into SVARs, since there will be an identity linking these assets, the interest rate, income and consumption. Also fiscal rules often involve the level of debt relative to some target value. To see the issues arising when allowing for stock variables in the context of a SVAR we look at a study by Cherif and Hasanov (2012). Cherif and Hasanov work with a SVAR involving four variables - the primary deficit to GDP ratio (pb_t) (public sector borrowing requirement), real GDP growth (Δy_t), the inflation rate of the GDP deflator (π_t) and the nominal average interest rate on debt (i_t). There is also a debt to GDP ratio d_t with d_{t-1} and d_{t-2} being taken to be "exogenous" regressors in all the structural equations. The SVAR is essentially recursive, except that in the first equation for pb_t the responses of pb_t to dy_t, π_t and i_t are set to .1, .07 and 0 respectively. The arguments for these values follow from the type of argument used by Blanchard and Perotti. This leaves us with the following (A, B) matrices

$$A = \begin{bmatrix} 1 & .1 & .07 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}, B = \begin{bmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix}.$$

Now to compute impulse responses to tax and spending shocks an allowance is needed for the fact that these will change the future path of d_t via a secondary impact on pb_t owing to its dependence on d_{t-1} and d_{t-2}. One can treat d_{t-1} and d_{t-2} as pre-determined for estimation purposes, but not when computing impulse responses for periods after the contemporaneous impact. To handle this Cherif and Hasanov add the identity

$$d_t = \frac{1 + i_t}{(1 + \pi_t)(1 + \Delta y_t)} d_{t-1} + pb_t$$

to the system and then solve for impulse responses from the augmented system. Because of the non-linearity, the impulses will depend on the values of pb_t, i_t etc. and the computations are therefore non-standard for EViews. Consequently we keep it within the relatively simple linear structure of EViews by replacing the debt equation with a log-linearized version.[10]

To derive this version let D_t be nominal debt, Y_t be real GDP, P_t be the price level and PB_t be the nominal primary deficit. Then the nominal debt identity is

$$D_t = (1 + i_t)D_{t-1} + PB_t.$$

Dividing this by P_tY_t produces

$$\frac{D_t}{P_tY_t} = (1 + i_t)\frac{D_{t-1}}{P_tY_t} + \frac{PB_t}{P_tY_t}$$

[10]If one wanted to use this identity it would be necessary to set the SVAR augmented with the identity to be a SYSTEM object, as described in Chapter 3.

Now the debt to GDP ratio (d_t) is $\frac{D_t}{P_tY_t}$, while the primary deficit to GDP ratio will be pb_t. Hence the debt equation is

$$\begin{aligned} d_t &= (1+i_t)\frac{D_{t-1}}{P_{t-1}Y_{t-1}}\frac{P_{t-1}Y_{t-1}}{P_tY_t}+\frac{PB_t}{P_tY_t} \\ &= (1+i_t)d_{t-1}\frac{P_{t-1}Y_{t-1}}{P_tY_t}+pb_t \\ &= (1+i_t)(1-\Delta p_t)(1-\Delta y_t)d_{t-1}+pb_t. \end{aligned}$$

This equation can be log-linearized by writing $d_t = d^*e^{\hat{d}_t}$, where d^* is the steady state value and $\hat{d}_t$ is the log deviation of d_t from that. Thus

$$\begin{aligned} d^*e^{\hat{d}_t} &= (1+i_t)(1-\Delta p_t)(1-\Delta y_t)d^*e^{\hat{d}_{t-1}}+pb_t \\ e^{\hat{d}_t} &= (1+i_t)(1-\Delta p_t)(1-\Delta y_t)e^{\hat{d}_{t-1}}+\frac{pb_t}{d^*}. \end{aligned}$$

Now, using $e^{\hat{d}_t} \simeq (1+\hat{d}_t)$,

$$1+\hat{d}_t = (1+\hat{d}_{t-1})(1+i_t)(1-\Delta p_t)(1-\Delta y_t)+\frac{pb_t}{d^*}.$$

Neglecting cross product terms this becomes

$$\hat{d}_t = \hat{d}_{t-1}+i_t-\Delta p_t-\Delta y_t+\frac{pb_t}{d^*}. \tag{4.25}$$

Since the identity does not affect estimation it can be done in a standard way in EViews using the A and B given above. The problem comes in computing the responses of variables to shocks. A first problem is that if we add d_t to the SVAR then one cannot have d_{t-1} and d_{t-2} treated as exogenous since d_{t-1} will necessarily be introduced into a SVAR simply through assigning some lag order to it. So we need to drop them from the exogenous variable set. A second issue is that there would be nothing to ensure that the debt accumulates according to the debt identity, i.e. if we form $\hat{d}_t$ from the data it would rarely equal the observed data $d_t - d^*$ owing to the linear approximation. A third issue is the problem that the identity does not have a shock attached to it. To get around the last two difficulties we construct a series $\tilde{d}_t$ as $\hat{d}_t$*+nrnd*.00001,* where the last term adds a small random number on to $\hat{d}_t$.

To understand how this works suppose that a SVAR(1) was fitted with the series $pb, dy, dp, in, \tilde{d}_t$. This setup doesn't fully capture Cherif and Hasanov since it only adds $\hat{d}_{t-1}$ on to the equations and not $\hat{d}_{t-2}$, but it is useful to start with the SVAR(1). Then this new system will have (A, B) as[11]

$$A = \begin{bmatrix} 1 & .1 & .07 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ -(1/d^*) & 1 & 1 & -1 & 1 \end{bmatrix}, B = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix}.$$

[11] This SVAR can be estimated with EViews using workfile *debt.wf1*.

Fitting the SVAR(1) will mean that the last equation involves a regression of $\phi_t = dhat - (pb_t/d^*) - in_t + \Delta p_t + \Delta y_t$ on $\tilde{d}_{t-1}, pb_{t-1}, in_{t-1}, \Delta p_{t-1}$ and Δy_{t-1}. The estimated equation coefficients using $\tilde{d}_t$ and $\hat{d}_t$ are the same to four decimal places so the debt accumulation identity effectively holds. The error introduced by using $\tilde{d}_t$ is very small. Hence the impulse responses of shocks on to the debt level can now be computed directly using the SVAR routine in EViews.

Now suppose that we want to capture the second lag of debt in the structural equations. This can be done by fitting an SVAR(2) with the same A and B above. Then the last equation would involve the regression of ϕ_t on $\{\tilde{d}_{t-1}, pb_{t-1}, in_{t-1}, \Delta p_{t-1},\ \Delta y_{t-1}\}$ and $\{\tilde{d}_{t-2}, pb_{t-2}, in_{t-2}, \Delta p_{t-2},\ \Delta y_{t-2}\}$. Using the data we find

$$\begin{aligned}
\phi_t &= .52\tilde{d}_{t-1} + .48\tilde{d}_{t-2} + .72pb_{t-1} - .48\Delta p_{t-1} - .48\Delta y_{t-1} + .48i_{t-1} \\
&= .52\tilde{d}_{t-1} + .48\tilde{d}_{t-2} + .48(\frac{pb_{t-1}}{d^*} - \Delta y_{t-1} - \Delta p_{t-1} + i_{t-1}) \\
&= \tilde{d}_{t-1} - .48\Delta\tilde{d}_{t-1} + .48\Delta\tilde{d}_{t-1} \\
&= \tilde{d}_{t-1} \\
&\Longrightarrow \tilde{d}_t = \tilde{d}_{t-1} + (pb_t/d^*) + in_t - \Delta p_t - \Delta y_t
\end{aligned}$$

as required. Fitting the SVAR(2) in the five variables produces the impulse responses for debt ratio shown in Figure 4.22.[12] It should be noted that there can be some problems with starting values and to change these one needs to use the sequence of commands after fitting the VAR(2) : ***Proc→Estimate Structural Factorization⟶Optimization Control → Starting values: Draw from Standard Normal.*** It is clear from the impulse response functions that it takes a long time for the debt to GDP ratio to stabilize after a shock.[13]

It is worth examining a comment by Cherif and Hasanov (2012, p 7) that "Similarly to Favero and Giavazzi (2007), we find that it is the change in debt that affects VAR dynamics as the coefficients on lagged debt are similar in absolute values but are of the opposite signs". If this were correct the evolution of debt would need to be described by

$$\hat{d}_t = \hat{d}_{t-1} + \phi_1^i \Delta\hat{d}_{t-1} - \phi_1^{\Delta p}\Delta\hat{d}_{t-1} - \phi_1^{\Delta y}\Delta\hat{d}_{t-1} + \phi_1^{pb}\hat{d}_{t-1} + ...,$$

i.e. the equation would be linear in $\Delta\hat{d}_t$, and so there would be no steady state debt to GDP ratio. It seems however that Cherif and Hasanov did not put the change in the debt into each of the SVAR equations, leaving the parameters attached to the levels variables $\hat{d}_{t-1}$ and $\hat{d}_{t-2}$ unconstrained. Nevertheless, the

[12] See also the code in *cherif_hasanov.prg*.

[13] If one wanted to allow for two lags in debt and only a single one in the other variables then one cannot use the SVAR option. It is necessary then to use the SYSTEM object to set up a system that can then be estimated. An example is given in the next application. Alternatively one can estimate the system by doing IV on each of the structural equations, and this is done in *debtsvar1iv.prg*. It is then necessary to compute impulse responses and this is done in *cherhas.prg*.

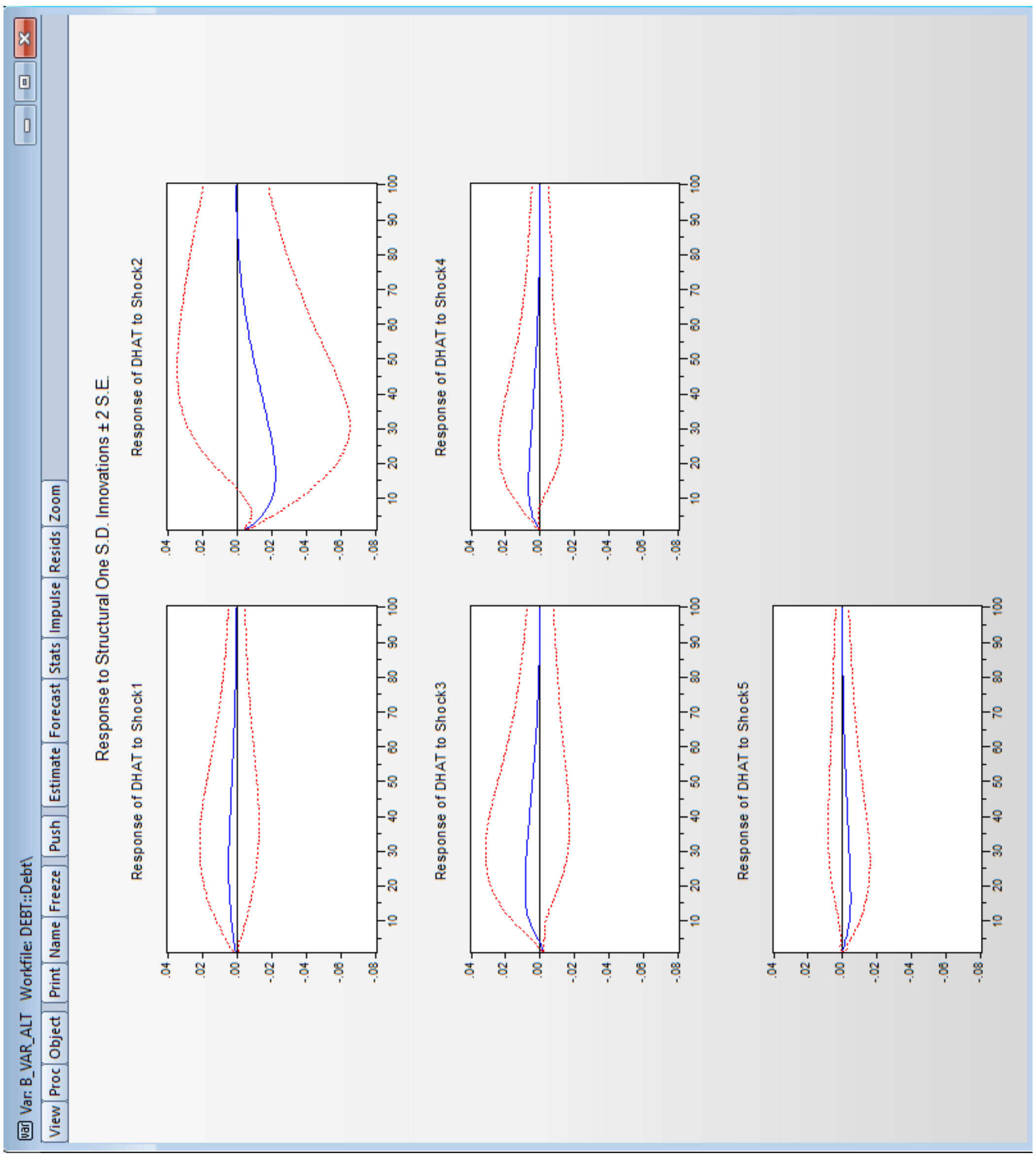


Figure 4.22: Impulse Responses for the SVAR(2) Model Incorporating the Debt Identity

parameters on these are such that they are very close to being equal and opposite and it is that which causes the very slow convergence.

4.6.5 Treating Exogenous Variables in an SVAR - the SVARX Model

Consider a 3 variable SVAR(1) where x_t is exogenous and z_{1t}, z_{2t} are endogenous. The exogeneity can be handled in two ways. One way involves including x_t, x_{t-1} as variables in each structural equation. In this case the first equation will be

$$\begin{aligned} z_{1t} &= a_{12}^0 z_{2t} + \gamma_1^0 x_t + \gamma_1^1 x_{t-1} + a_{11}^1 z_{1t-1} \\ &\quad + a_{12}^1 z_{2t-1} + \varepsilon_{1t}, \end{aligned}$$

and equations like this constitute an SVARX system. EViews can handle this by using the exogeneity option when describing the VAR specification. There are however some specifications for exogenous variables that cannot be handled with the exogeneity option. One of these arises if it is undesirable to have x_t enter into *every* structural equation. Another would be if one wanted to compute either the dynamic multipliers with respect to x_t or to shocks into x_t, where the shock is defined by some SVAR. Lastly, in the case of a small open economy, where the z_t would be domestic variables and x_t foreign variables, then we would want to ensure that no lagged values of the z_t impact upon the x_t. This rules out the possibility of fitting an SVAR in z_t and x_t with standard EViews software.

It would be possible to suppress the dependence of z_t on x_t by inserting zeroes into the specification of the A_0 matrix but there is no simple way of setting the similar elements in A_j $(j > 1)$ to zero. Because of this z_{t-j} would affect x_t. Of course it may be that the coefficients in A_1 etc. are small but, with a large number to be estimated, it seems more sensible to constrain them to be zero. To deal with these cases one needs to create a SYSTEM object, just as was done with the restricted Brazilian VAR in Chapter 2.

To illustrate the method we examine impulse responses for a Brazilian SVAR when the EViews exogeneity option is used, and then describe how to compute either dynamic multipliers or impulse responses to shocks for the exogenous variables. Because Brazil is taken to be a small open economy we do not want domestic variables to impact upon the foreign variables. The workfile is *brazil.wf1*. The SVAR is formulated in terms of five domestic variables n_t (GNE), y_t (GDP), $infl$ (inflation), int (interest rate) and rer (real exchange rate). The presence of both n_t and y_t is because Brazil is an open economy and so the first of these captures aggregate demand while the second is supply. As mentioned in Chapter 2 these variables are measured as a deviation from permanent components. The exogenous variables are taken to be two world variables - $ystar_t$(external output) and rus_t (an external interest rate). A VAR(1) is fitted owing to the short data set. Figure 4.23 shows the required EViews commands.

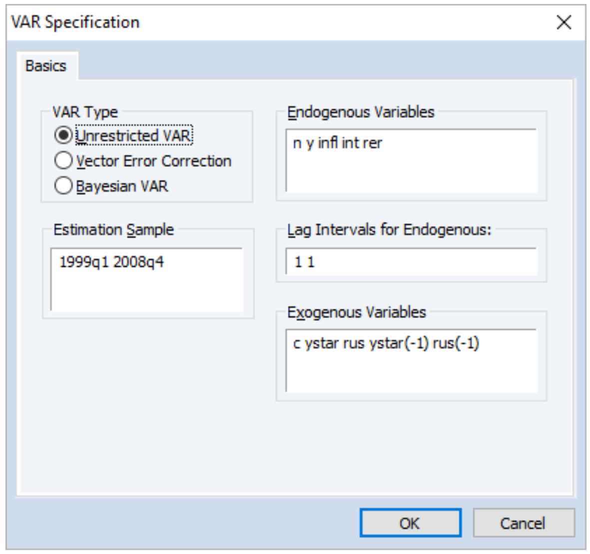


Figure 4.23: The Brazilian SVAR(1) Model with the Foreign Variables Treated as Exogenous

The exogenous variables have been entered in this way so as to get a system representation that can be modified relatively easily.

Once this VAR is run, invoking the commands from the EViews pull-down menus of ***Proc*** $\rightarrow$ ***Make System*** $\rightarrow$ ***Order by Variable*** produces equivalent model code that can be edited to insert structural equations for *ystar* and *rus* that do not allow for any feedback between the domestic and foreign sectors. We set these up as in Figure (4.24).

The reason for the structure chosen is that by setting the coefficients $C(53)...C(57)$ equal to zero, the responses of the domestic variables to the foreign variables will be dynamic multipliers, whereas, if all (i.e., 57) of the original coefficients are estimated then non-zero impulse responses can be found for shocks to the foreign variables. Note that there is now is no influence of the domestic variables on the foreign ones.

This system can be estimated by OLS. Then the estimated coefficients $C(\cdot)$ can be mapped into A_0 and A_1 using *brazsvarbig.prg.* This program also produces impulse responses (and dynamic multipliers if the coefficients $C(53) - C(57)$ are not estimated). Figure 4.25 shows the impulse responses to the SVAR using the exogeneity specification. The real exchange rate is measured so that a rise represents an appreciation. We note that the effect of interest rates upon demand (n_t) is substantial but that on supply (y_t) is small (as it should be). A demand shock raises inflation, while a positive shock to supply (the y shock) reduces it (Figure 4.26). The exchange rate appreciates in response to an interest rate shock and also to a foreign output shock, given the positive parameter estimates for YSTAR on RER in the VAR. (Figure 4.27).

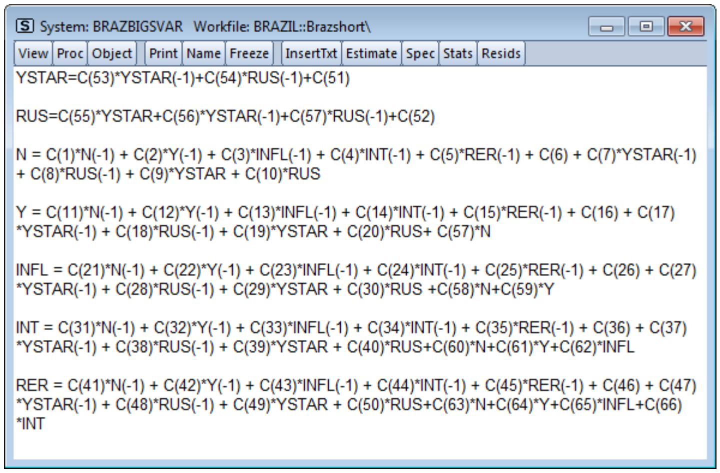


Figure 4.24: The Brazilian SVAR(1) Model Absent Lagged Feedback Between the Foreign and Domestic Sectors

An appreciation leads to a rise in demand and this can reflect the lower prices for commodities. There is a negative response by domestic output to a real exchange rate appreciation (Figure 4.28).

We can also compare the impulse responses that would be found if we treated all the seven variables (both domestic and foreign) as endogenous and fitted a recursive SVAR(1) using EViews with that in which there are no lagged values of the domestic variables impacting on the foreign sector, i.e. the model is that in the SYSTEM object BRAZBIGSVAR. Because the shocks for the foreign variables will be different with each specification (there are more regressors in the y_t^* equation when lagged feedback is allowed) we set the standard deviation of the shock to that under the feedback solution. Then the impulse responses of the real exchange rate to a foreign output shock for each of the systems are in Figure (4.29). There is clearly not a great deal of difference between the responses at short horizons.

4.6.6 Restrictions on Parameters and Partial Exogeneity: External Instruments

A literature has emerged where it is possible to estimate the SVAR by using "external instruments". These variables function like instruments in that they are uncorrelated with some of the structural shocks and therefore exogenous to the corresponding structural equations. However, they are correlated with other structural shocks, i.e. there is only partial exogeneity. Applications have been made by Olea *et al.* (2013) and Mertens and Ravn (2012). In the latter the external instrument is a set of "narrative" fiscal shocks constructed by Romer and Romer (2010), while Olea *et al.* use a variable constructed by Kilian (2008)

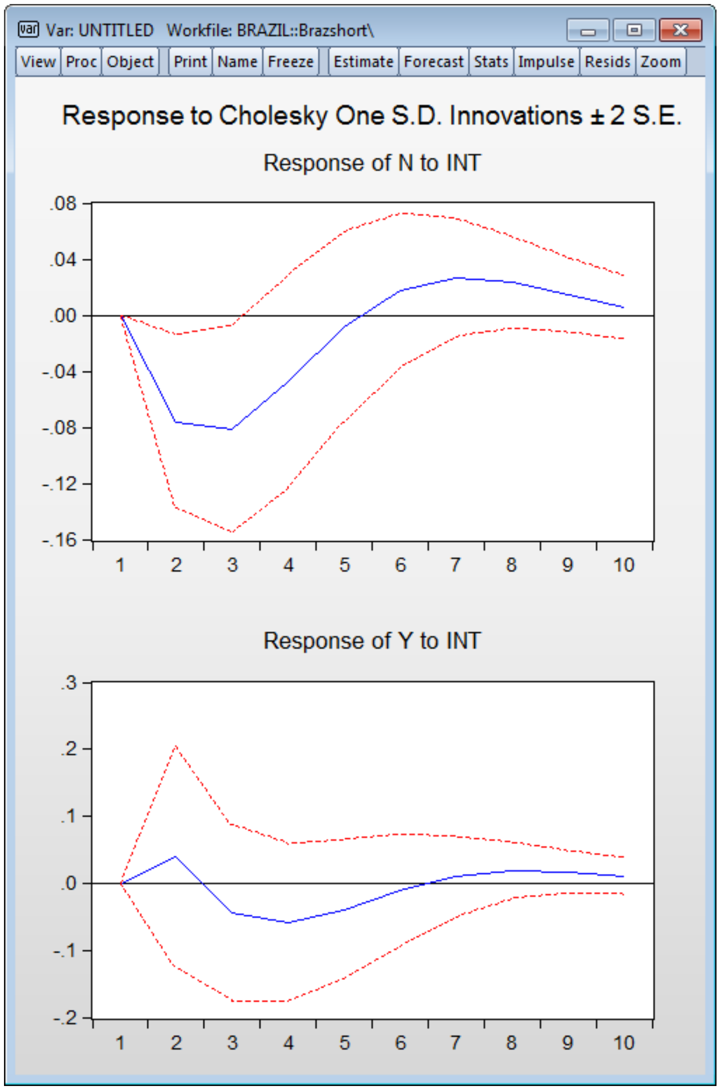


Figure 4.25: Response of Demand (N) and Income (Y) to Interest Rates for the Brazilian SVAR With Foreign Variables Exogenous

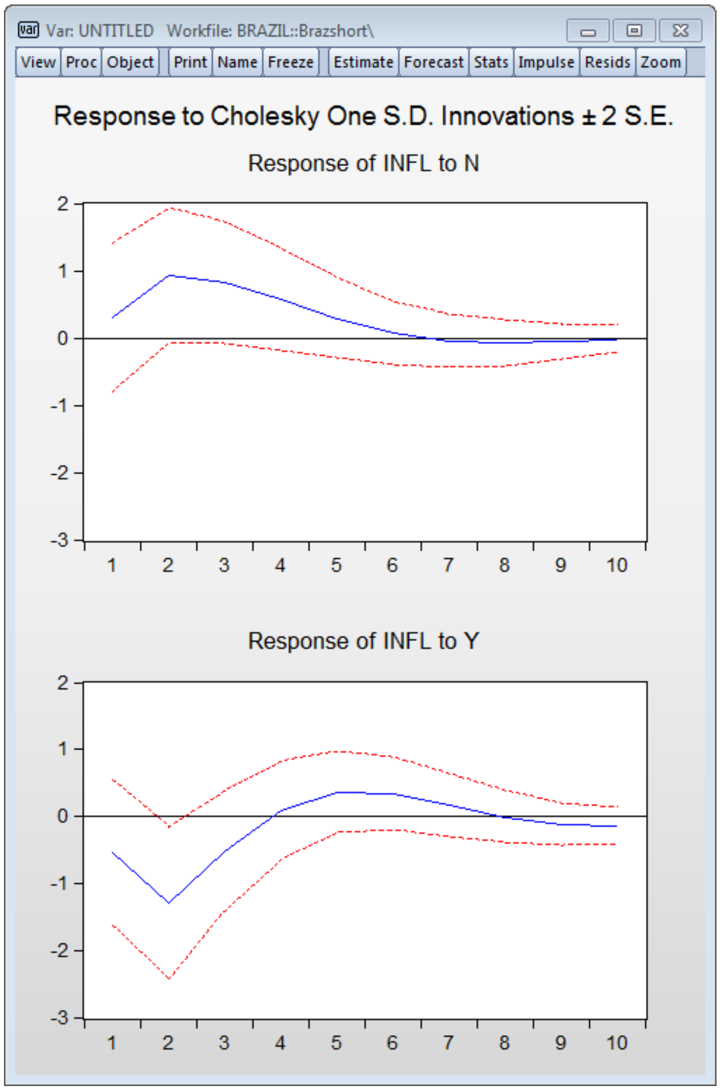


Figure 4.26: Response of Inflation (INFL) to Demand (N) and Output Shocks (Y) for the Brazilian SVAR With Foreign Variables Exogenous

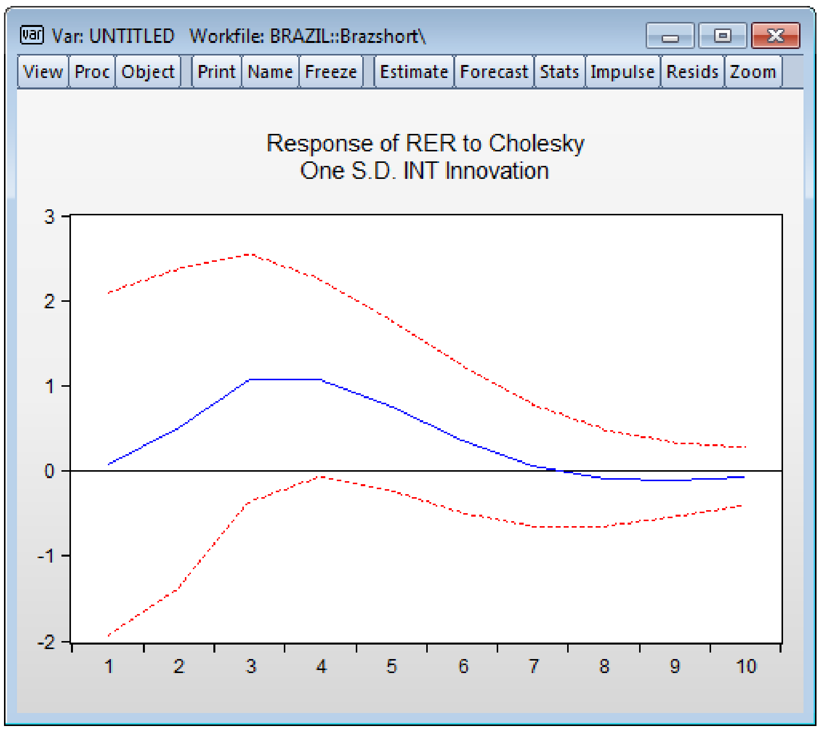


Figure 4.27: Response of the Real Exchange Rate (RER) to an Interest Rate (INT) Shock for the Brazilian SVAR With Foreign Variables Exogenous

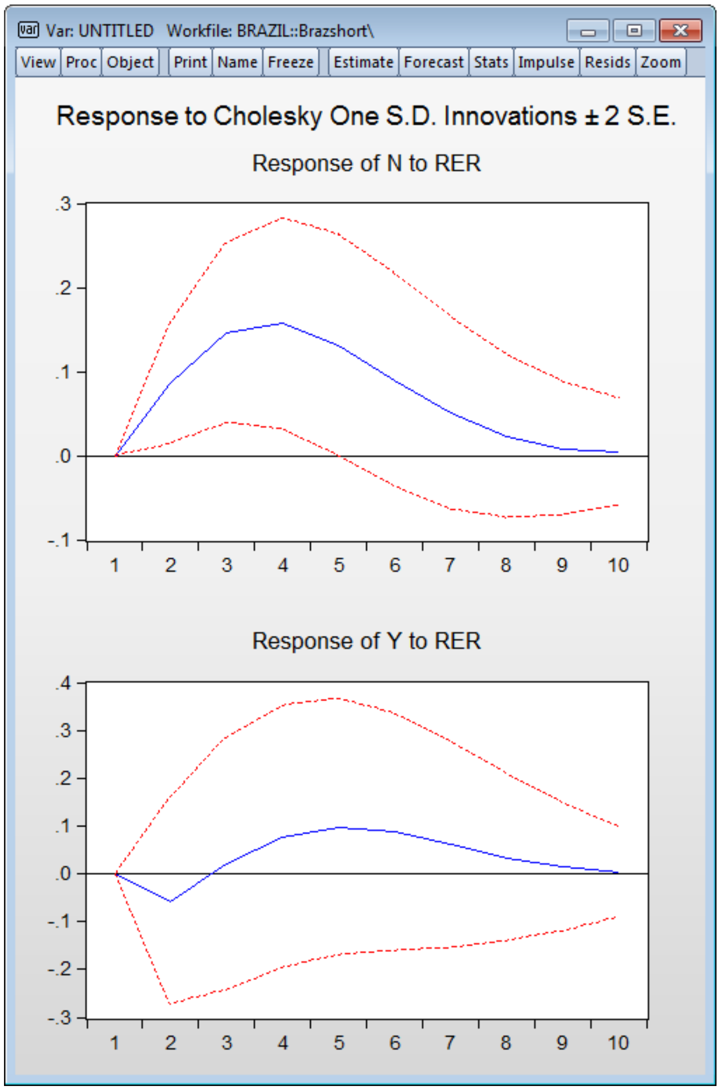


Figure 4.28: Response of Demand (N) and Output (Y) to the Real Exchange Rate (RER) for the Brazilian SVAR With Foreign Variables Exogenous

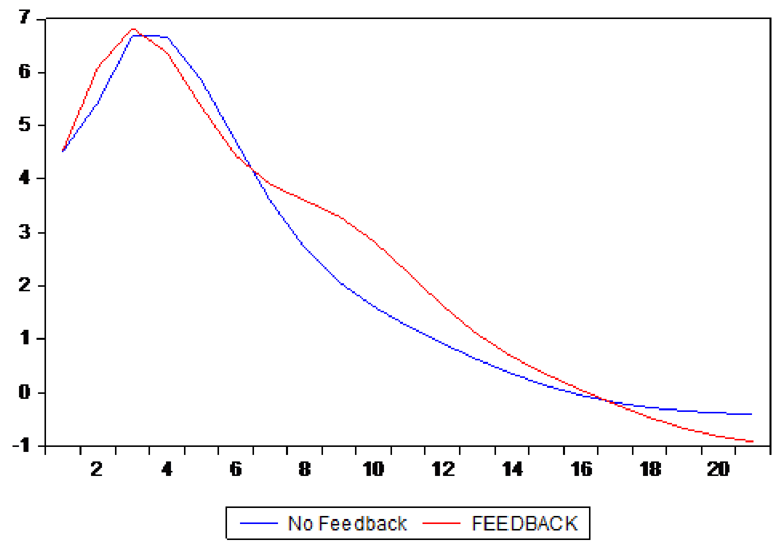


Figure 4.29: Response of Real Exchange Rate (RER) to a Foreign Output Shock (YSTAR) in the Brazilian SVAR Models

on the oil supply shortfall. To use these instruments effectively it needs to be ensured that they do not appear in the structural equations whose shocks they are uncorrelated with. If they did appear moment conditions would be "used up" in estimating the coefficients of such variables in each structural equation in the model. Hence there must be some parameter restrictions in the system of equations of the larger system that incorporates the instruments.

To see an application of the methodology and how it might be implemented in EViews 9 we follow Mertens and Ravn (2012) and return to the Blanchard and Perotti model. In its general form it is

$$\begin{aligned} z_{1t} &= a_1 z_{3t} + a_2' \varepsilon_{2t} + lags + \varepsilon_{1t} \\ z_{2t} &= b_1 z_{3t} + b_2' \varepsilon_{1t} + lags + \varepsilon_{2t} \\ z_{3t} &= \delta_1 z_{1t} + \delta_2 z_{2t} + lags + \varepsilon_{3t} \end{aligned} \tag{4.26}$$

Now they set $b_1 = 0$ which seems unexceptional. Because this system could not be estimated they reduced the number of unknown parameters to six by first fixing a_1 to 2.08, thereafter setting either a_2' or b_2' to zero, leaving only one of them to be estimated. Mertens and Ravn propose estimating the remaining eight parameters of the system by using external instruments (or what they refer to as a proxy). Given that there is such an instrument, m_t, it is assumed that $E(m_t \varepsilon_{1t}) \neq 0$, $E(m_t \varepsilon_{2t}) = 0$ and $E(m_t \varepsilon_{3t}) = 0$.

To see how this might be handled in EViews, we augment the SVAR used by Blanchard and Perotti with an equation for m_t. Then the SVAR system becomes

$$z_{1t} = a_1 z_{3t} + a_2' \varepsilon_{2t} + lags + \varepsilon_{1t} \tag{4.27}$$
$$z_{2t} = b_2' \varepsilon_{1t} + lags + \varepsilon_{2t} \tag{4.28}$$
$$z_{3t} = \delta_1 z_{1t} + \delta_2 z_{2t} + lags + \varepsilon_{3t} \tag{4.29}$$
$$m_t = lags + \rho \varepsilon_{1t} + \varepsilon_{mt}, \tag{4.30}$$

where $E(\varepsilon_{mt}\varepsilon_{jt}) = 0, j = 1, ..., 3$.

This structure incorporates the restrictions pertaining to the external instrument m_t given above. Then (4.27) - (4.30) is a standard SVAR with ten restrictions coming from the fact that all shocks are uncorrelated. Two of these are needed to estimate ρ and the standard deviation of ε_{mt}, while the remaining eight are used to estimate the parameters a_1, a_2' etc. The model is therefore exactly identified. To estimate it using EViews we use the (A, B) technology for Az_t=Bη_t+lags, with these matrices being defined by

$$A = \begin{bmatrix} 1 & 0 & -a_1 & 0 \\ 0 & 1 & 0 & 0 \\ -\delta_1 & -\delta_2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} \sigma_1 & a_2'\sigma_2 & 0 & 0 \\ b_2'\sigma_1 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ \rho & 0 & 0 & \sigma_m \end{bmatrix}.$$

We note that in the applications by Mertens and Ravn and Olea *et al.* there do not seem to be any lags in the equation for m_t. If one wants to impose such a specification then it would be necessary to impose zero restrictions upon the A_j (i.e., lag) matrices using the SYSTEM object in the same way as described in the preceding sub-section.

4.6.7 Factor Augmented SVARs

Often many variables may be available to a researcher which are expected to influence macro-economic outcomes. Thus financial factors and confidence might be important to household decisions. Because there is rarely a single measure of these factors, there is a tendency to utilize many approximate measures, particularly involving data surveying the attitudes of financial officers, households or businesses. There are far too many of these measures to put them all into a SVAR, so some aggregation is necessary. For a small system involving macroeconomic variables such as the unemployment rate, industrial production and employment growth, Sargent and Sims (1977) found that two dynamic factors could explain 80 percent or more of the variance of those variables. One of the factors was primarily associated with real variables and the other with inflation. Bernanke *et al.* (2005) extended this approach and they proposed augmenting a SVAR with a small number of factors.

We considered the Bernanke *et al.* (2005) Factor Augmented VAR model in Chapter 2. There are two difficulties in implementing a factor oriented approach. One is how to measure the factors and the other is how to enter these into a SVAR, particularly in deciding on how to estimate the contemporaneous part of the SVAR. In Chapter 2 following Bernanke *et al.* three factors were extracted from a set of 119 series X_t. These factors were the principal components and will be referred to as $\hat{F}_t$ (in Chapter 2 these were called *pc1_x, pc2_x* etc.). One variable not in X_t was R_t (the Federal Funds Rate). We might think about forming a SVAR with $\hat{F}_t$ and R_t present to capture the effects of monetary shocks upon the factors $\hat{F}_t$. However, because R_t would react to $\hat{F}_t$, and there is no reason to think that interest rates won't contemporaneously react to some of the variables in X_t from which the principal components are computed, then a recursive SVAR model with $\hat{F}_t$ and R_t would be inappropriate. So one needs to impose some other identification assumption upon the SVAR to enable it to be estimated.

The system Bernanke *et al.* have in mind consists of

$$X_t = \Lambda F_t + \Lambda^r R_t + e_t \tag{4.31}$$

$$F_t = \Phi_{11} F_{t-1} + \Phi_{12} R_{t-1} + \varepsilon_{1t} \tag{4.32}$$

$$R_t = \Phi_{21} F_{t-1} + \Phi_{22} R_{t-1} + B F_t + \varepsilon_{2t}, \tag{4.33}$$

where X_t is an $N \times 1$ vector of "informational variables", F_t is a $K \times 1$ vector of factors and R_t is a nominal interest rate. N is much greater than K. (4.31) is their equation (2), while (4.32) – (4.33) correspond to their equation (1), except it is written as a SVAR rather than a VAR. The SVAR structure comes from their examples, in which the factor F_t enters contemporaneously into the central bank's decision rule for interest rates along with the statement (p 401) that "all the factors entering (1) respond with a lag to changes in the monetary policy instrument". We will focus on the empirical part of the paper where there is a single observable factor - the interest rate - although they also suggest that R_t might be replaced by a vector Y_t of observables. In their application X_t in (4.31) is a large data set of 119 variables, where R_t is excluded from X_t.[14] This data set consists of "fast moving" and "slow moving" variables, where the difference is that the 70 slow moving variables X_t^s in X_t do not depend contemporaneously on R_t. The fast moving variables will be X_t^f and they do have a contemporaneous dependence.

Given the factor structure, the key identification assumption is Bernanke *et al.'s* suggestion that the slow moving variables depend contemporaneously on the factors but not the interest rate, i.e.

$$X_t^s = G F_t + v_t.$$

Now suppose that K principal components (PCs) are extracted from the N^s elements in X_t^s. Then Bai and Ng (2006, 2013) show that the asymptotic relation between the principal components (PC_t^s) and the K factors will be $F_t = (H \times PC_t^s) + \xi_t$, where ξ_t is $O_p(\frac{1}{\sqrt{N^s}})$. Provided $N^s \to \infty$ such that $\frac{\sqrt{N^s}}{T} \to 0$ the principal components asymptotically span the space of the factors. Bai and Ng (2013) consider what would be needed for H to be the identity

[14] Here we follow the Matlab program that Boivin supplied to reproduce their results.

matrix, and state some conditions that would need to be enforced in forming the principal components, but these methods are unlikely to have been used in the FAVAR applications. Therefore replacing F_t by $F_t = HPC_t^s$ in (4.32) and (4.33) will give $A_{11}^1 = H^{-1}\Phi_{11}H$ etc.

$$PC_t^s = A_{11}^1 PC_{t-1}^s + A_{12}^1 R_{t-1} + H^{-1}\epsilon_{1t} \tag{4.34}$$

$$R_t = A_{21}^1 PC_{t-1}^s + A_{22}^1 R_{t-1} + A_{21}^0 PC_t^s + \epsilon_{2t} \tag{4.35}$$

This is an SVAR in R_t and PC_t^s except that, unlike a regular SVAR, the shocks in the PC_t^s equations are contemporaneously correlated.[15] The presence of an unknown H in $H^{-1}\varepsilon_t$ will mean that it is not possible to calculate impulse responses with respect to the shocks ε_{1t}. It is also not possible to estimate H from the covariance matrix of $H^{-1}\varepsilon_{1t}$, because the former has K^2 elements and the latter has $K \times (K+1)/2$ elements. Nevertheless, once the SVAR in PC_t^s and R_t is estimated, the impulse responses to the monetary shock ε_{2t} can be found.

To determine the impact upon members of X_t^f and X_t^s it is necessary to express these in terms of the SVAR variables. Thus, using the mapping between factors and principal components,

$$X_t^s = GH(PC_t^s) + v_t$$

and the regression of X_t^s on PC_t^s consistently estimates GH, because F_t is assumed uncorrelated with v_t. Consequently, the impulse responses of X_t^s to the monetary shocks can be computed. In general

$$X_t = (\Lambda H)PC_t^s + \Lambda^r R_t + e_t, \tag{4.36}$$

and the same process gives the weights ΛH and Λ^r.

Now, the K principal components of X (PC_t^x) can be written as

$$PC_t^x = w' X_t,$$

where w' are the matrix of weights found from the PC analysis. Then, using (4.36),

$$\begin{aligned} PC_t^x &= w'[(\Lambda H)PC_t^s + \Lambda^r R_t + e_t] \\ &= (w'\Lambda H)PC_t^s + w'\Lambda^r R_t + w'e_t \\ &= G_1 PC_t^s + G_2 R_t + w'e_t, \end{aligned} \tag{4.37}$$

and a regression of PC_t^x against PC_t^s and R_t consistently estimates $w'\Lambda H$ and $w'\Lambda^r$. Adding this to the system of SVAR equations results in the complete system to find impulse responses.

[15] Because the same regressors appear in all the equations in (4.34) the OLS estimator of the parameters of those equations is efficient. But the non-diagonal covariance matrix for the errors coming from $H^{-1}\varepsilon_{1t}$ means that this would need to be allowed for when getting the standard errors of $H^{-1}\Phi_{11}H$ and $H^{-1}\Phi_{12}H$.

The above analysis describes a SVAR that can be used to find impulse responses of variables in X_t to interest rate shocks. However, this is not the system that Bernanke *et al.* work with. Rather they first regress PC_t^x against PC_t^s and R_t to estimate G_1 and G_2 in (4.37), and then use $\tilde{F}_t = PC_t^x - G_2 R_t$ in a block recursive SVAR ordered as $(\tilde{F}_t, R_t)$. Can one recover such a SVAR from the one involving PC_t^s and R_t? The answer is in the negative. To see why look at the definition of $\tilde{F}_t = PC_t^x - G_2 R_t$. From (4.37) that means $\tilde{F}_t = G_1 PC_t^s + w' e_t$. Hence, assuming that the number of principal components PC_t^s is not larger than PC_t^x, it follows that $PC_t^s = \Phi_1 \tilde{F}_t + \Phi_2 e_t$. Substituting this into (4.34) and (4.35) shows that the two equations will have e_t entering into the error terms of both equations i.e. $\tilde{F}_t$ will be correlated with the error term for the R_t equation.[16] Consequently, if one regresses R_t on $\tilde{F}_{t-1}, R_{t-1}$ and $\tilde{F}_t$ one will get inconsistent estimators for the parameters of the R_t equation, and therefore the same will be true of the impulse responses.[17] Hence the method Bernanke *et al.* use to account for simultaneity does not deliver consistent estimators of the impulse responses of interest. One needs to use the SVAR formulated with the slow-moving variables and then recover the impulse responses to X_t (say industrial production ip_t) from (4.36).

In practice some further adjustments are also required to compute the impulse responses of interest. First, since variables like ip_t have been standardized, it is necessary to multiply by the standard deviation of ip_t to get back to the responses of the original industrial production variable. Second, a variable such as industrial production enters into X_t in growth form (specifically log difference form). Therefore, to get the impact on the *levels* of industrial production, it is necessary to accumulate the impulse responses. Lastly, it is necessary to form the exponential of these in order to arrive at the responses of the level of industrial production.

Figure 4.27 presents some impulse responses for a range of variables to a shock in the Federal Funds Rate. This example is taken from Bernanke *et al.* The two impulses presented in each graph are for the FAVAR based on just the slow moving variables and also the results from the "purging" method used by Bernanke *et al.* The size of the shock is the same as used by those authors. The variables presented are LEHCC (Average Hourly Earnings of Construction Workers), CMCDQ (Real Personal Consumption Expenditures), PUNEW (the CPI, all items), EXRJAN (Yen/ Dollar Exchange Rate), IP (Industrial Production Index) and HSFR (Total Housing Starts). For all variables except housing starts the impulses are accumulated, since those variables were measured as the change in the logarithms. Consequently, the responses measure the impact of interest rate shocks upon the level of the CPI, industrial production etc. As Fisher *et al.* (2015) point out this specification means that the level of industrial production and consumption will be permanently affected by a one period interest rate shock, and this is apparent from the graphs. There are differ-

[16] It is also the case that e_{t-1} enters into both error terms and so the system will not be a SVAR but a SVARMA process.

[17] Boivin and and Giannoni (2009) suggest an iterated version of this strategy but it also fails to consistently estimate the parameters of the SVAR being used.

ences between the two sets of responses, notably for industrial production, the exchange rate and the CPI. The inconsistent estimates found from using the Bernanke *et al.* approach are much larger than those found using the SVAR in slow moving variables.[18]

Quite a few applications of FAVAR models exist. Eickmeier and Hofmann (2013) use a (FAVAR) estimated to analyze the US monetary transmission via private sector balance sheets, credit risk spreads, and asset markets, and to study the "imbalances" observed prior to the global financial crisis - high house price inflation, strong private debt growth and low credit risk spreads. Lombardi *et al.* (2010) using a set of non-energy commodity price series extract two factors and place these in a VAR together with selected macroeconomic variables.

4.6.8 Global SVARs (SGVARs)

One example of a VARX system is the Global VAR (GVAR). In this there is a typical VAR equation for the $i'th$ country which expresses z_{it} (say the log of GDP for the i'th equation) as a function of a global variable z^*_{it}

$$z_{it} = A_i z_{t-1} + \delta_i z^*_{it} + \varepsilon_{it}.$$

Here $z^*_{it} = \sum_{j=1,j\neq i}^{n} \omega_{ij} z_{it}$ is a "world" variable from the perspective of the i'th country, ω_{ij} are trade or financial flow weights. This means that the value of z_{it} for the $i'th$ country value does not appear in z^*_{it}. GVARS mostly use generalized impulse responses and so are not really SVARs, but recently some SGVARs have been proposed. We look briefly at this literature as one has to be careful using it, and that has not been true of some applications.

The example we will work with is a SGVAR with 3 countries where, for simplicity, lags will be ignored. This will yield the three equations

$$\begin{aligned} z_{1t} &= \delta_1 z^*_{1t} + \varepsilon_{1t} = \delta_1(\omega_{12} z_{2t} + \omega_{13} z_{3t}) + \varepsilon_{1t} \\ z_{2t} &= \delta_2 z^*_{2t} + \varepsilon_{2t} = \delta_2(\omega_{21} z_{1t} + \omega_{23} z_{3t}) + \varepsilon_{2t} \\ z_{3t} &= \varepsilon_{3t}, \end{aligned} \tag{4.38}$$

where ε_{jt} are structural shocks and the third country is the "numeraire" in the sense that it has no corresponding z^*_{3t}. This is an SVARX due to z^*_{it} being in the equations and being treated as exogenous.

Can we estimate the first country equation with OLS? To answer this we need to look at the correlation between z^*_{1t} and ε_{1t}, which is found from

$$\begin{aligned} E(z^*_{1t}\varepsilon_{1t}) &= E(\omega_{12} z_{2t}\varepsilon_{1t} + \omega_{13} z_{3t}\varepsilon_{1t}) \\ &= E(\omega_{12} z_{2t}\varepsilon_{1t}) \\ &= \omega_{12}\delta_2\omega_{21} E(z_{1t}\varepsilon_{1t}). \end{aligned}$$

[18]The EViews code for estimating the original FAVAR model by Bernanke *et al.* can be found in the sub-directory "BBE" in the "EViews Content" folder. See the files named *bbe_f1.prg* and *bbe_f2.prg.* The program *bbe_f1_alt.prg* implements the alternative approach described in this section.

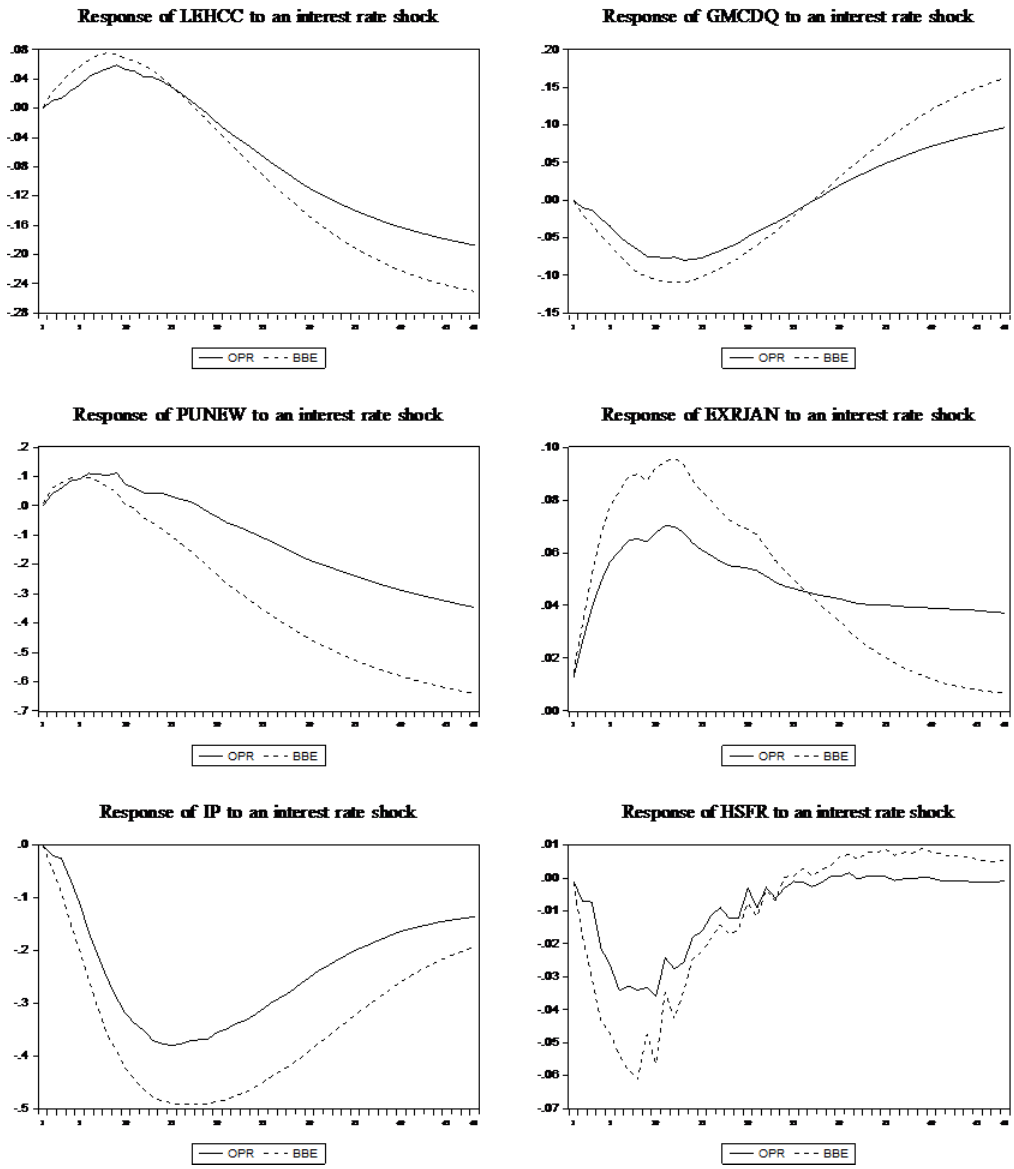


Figure 4.30: Comparison of Impulse Responses to Interest Rate Shocks from OPR and Bernanke *et al.*

Clearly OLS is not consistent unless either $\omega_{12} = 0, \omega_{21} = 0$ (unlikely) or $\delta_2 = 0$ (unlikely as well since it would mean that for the second country there are no foreign influences). It would be different if (4.38) had the form

$$z_{2t} = \delta_{21}\omega_{21}z_{1t} + \delta_{23}\omega_{23}z_{3t} + \varepsilon_{2t},$$

that is if $\delta_{21} \neq \delta_{23}$. Then $\delta_{21} = 0$ might be imposed and the system would be a recursive one with ordering (z_{3t}, z_{2t}, z_{1t}). But the SGVAR imposes $\delta_{21} = \delta_{23}$ and, although this is parsimonious, it makes exogeneity of z^*_{1t} in the z_{1t} equation implausible.

A number of applications have suggested that one can consistently estimate the parameters in the equations of this SGVARX. However, what is being estimated is not the first structural equation but a conditional equation describing $E(z_{1t}|z^*_{1t})$ viz.

$$E(z_{1t}|z^*_{1t}) = \delta_1 z^*_{1t} + E(\varepsilon_{1t}|z^*_{1t}).$$

Under joint normality of the shocks $E(\varepsilon_{1t}|z^*_{1t}) = \rho z^*_{1t}$, where ρ is proportional to $\mathrm{corr}(\varepsilon_{1t},\, z^*_{1t})$, and this is not zero unless $\delta_2 = 0$. Hence the *conditional equation* is

$$E(z_{1t}|z^*_{1t}) = (\delta_1 + \rho)z^*_{1t},$$

and so what is being consistently estimated is $(\delta_1 + \rho)$ rather than δ_1, which was the coefficient of the basic SGVARX model. For some purposes, e.g. forecasting, it may be irrelevant that what is being estimated is $\delta_1 + \rho$ rather than δ_1, but this not true for impulse responses.

EViews does not enable researchers to easily estimate GVAR models. There is software using MATLAB that is available at

$$http://www-cfap.jbs.cam.ac.uk/research/gvartoolbox/download.html$$

4.6.9 DSGE Models and the Origins of SVARs

Theoretical models like DSGE have structural equations such as

$$z_{1t} = \phi z_{1t-1} + \psi E_t z_{1t+1} + \rho z_{2t} + v_{1t}. \tag{4.39}$$

The system they are part of generally reduces to a VAR(1) in the variables provided there is no serial correlation in the structural shocks. If shocks are first order serially correlated then the VAR for the system has a maximum order of two. The question then is what does this imply about SVARs?

Suppose there is a 3 variable structural system containing the equation above and there is a VAR(1) solution to it. Then this implies

$$z_{1t} = b^1_{11}z_{1t-1} + b^1_{12}z_{2t-1} + b^1_{13}z_{3t-1} + e_{1t}$$

and

$$E_t z_{1t+1} = b^1_{11}z_{1t} + b^1_{12}z_{2t} + b^1_{13}z_{3t}. \tag{4.40}$$

Eliminating expectations in (4.39) using (4.40) it becomes

$$z_{1t} = \phi z_{1t-1} + \psi(b_{11}^1 z_{1t} + b_{12}^1 z_{2t} + b_{13}^1 z_{3t}) + \rho z_{2t} + v_{1t}$$

which can be written as

$$(1 - \psi b_{11}^1) z_{1t} = \phi z_{1t-1} + (\psi b_{12}^1 + \rho) z_{2t} + \psi b_{13}^1 z_{3t} + v_{1t}$$

Gathering terms we get

$$\begin{aligned} z_{1t} &= a_{11}^1 z_{1t-1} + a_{12}^0 z_{2t} + a_{13}^0 z_{3t} + \varepsilon_{1t} \\ a_{11}^1 &= \frac{\phi}{(1-\psi b_{11}^1)}, a_{12}^0 = \frac{(\psi b_{12}^1 + \rho)}{(1-\psi b_{11}^1)} \\ a_{13}^0 &= \frac{\psi b_{13}^1}{(1-\psi b_{11}^1)}, \varepsilon_t = \frac{v_{1t}}{(1-\psi b_{11}^1)} \end{aligned} \tag{4.41}$$

This equation is a structural equation in an SVAR. There are three parameters in this SVAR equation and three parameters in the original DSGE equation. Hence the SVAR equation is just a re-parameterization of the DSGE one. Note the difference to a standard SVAR - z_{2t-1}, z_{3t-1} are *excluded* from the SVAR. Thus DSGE models certainly employ *exclusion restrictions* in estimation in order to get identification. In practice other restrictions are also used by DSGE models to get identification and these arise from the fact that there are a smaller number of unknown parameters in the DSGE model than the implied SVAR model, i.e. the parameters b_{ij}^1 may be functions of less than three parameters. Hence there are restrictions upon the SVAR parameters in (4.41). These are often referred to as cross-equation restrictions since the fundamental set of parameters appear in other equations as well. Pagan and Robinson (2016) have an extended discussion of the relationship between DSGE and SVAR models.

4.7 Standard Errors for Structural Impulse Responses.

Because $\hat{C}_j = \hat{D}_j \hat{C}_0$ the impulse responses are now combinations of those found with the VAR. We discussed how to find standard errors for D_j in Chapter 3, but it is now apparent that there is a complication when evaluating those of $\hat{C}_j$. This comes from the fact that it is a product of “two” random variables $\hat{C}_0$ and $\hat{D}_j$. Normally asymptotic standard errors for such products are found by using the delta method, but that assumes both $\hat{C}_0$ and $\hat{D}_j$ can be regarded as normally distributed in large samples. It may be a reasonable assumption for $\hat{D}_j$ but it is far less likely to be true for $\hat{C}_0$, and we now examine why this might be so.

To see the argument in its simplest form consider a two variable structural system consisting of a money demand function and an interest rate rule. As it is C_0 which is of interest it will be assumed that there are no lagged values in the equations and that income effects are set to zero (if income is introduced one would need a three variable system). Then the normalized system is

$$m_t = a_{12}^0 i_t + \varepsilon_{1t} \tag{4.42}$$
$$i_t = a_{21}^0 m_t + \varepsilon_{2t}. \tag{4.43}$$

In matrix form this becomes

$$\begin{bmatrix} 1 & -a_{12}^0 \\ -a_{21}^0 & 1 \end{bmatrix} \begin{bmatrix} m_t \\ i_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

and the contemporaneous impulse responses will be

$$C_0 = A_0^{-1} = \begin{bmatrix} 1 & -a_{12}^0 \\ -a_{21}^0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{1-a_{12}^0 a_{21}^0} & \frac{a_{12}^0}{1-a_{12}^0 a_{21}^0} \\ \frac{a_{21}^0}{1-a_{12}^0 a_{21}^0} & \frac{1}{1-a_{12}^0 a_{21}^0} \end{bmatrix}.$$

Notice the estimated impulse responses are functions of the estimators of the structural coefficients a_{12}^0 and a_{21}^0. Consequently, two things could go wrong. One is that $\hat{a}_{12}^0$ and $\hat{a}_{21}^0$ may be non-normal (in finite samples). As emphasized before these are effectively estimated by instrumental variables and we know that, if the instruments are weak, then these distributions can be far from normal, even in quite large sample sizes. If the system is recursive then that should not be an issue, as the instruments are the variables themselves, but in other instances one cannot be so confident. The other problem comes from the fact that the contemporaneous impulse responses C_0 involves ratios of random variables. Whilst products of random variables can generally be handled quite well with the δ-method, this is not so true when ratios are involved. Figure 4.31, taken from Pagan and Robertson (1998), shows an example of this using the money supply and demand model in Gordon and Leeper (1994), which is a more sophisticated version of the two equation system discussed above. It is clear that normality does not hold for the estimators of the coefficients a_{ij}^0 and this goes over to affect the impact of a money supply shock upon the interest rate ($\eta = \frac{a_{21}^0}{1-a_{12}^0 a_{21}^0}$).

One needs to be cautious with the "confidence intervals" coming out of packages such as EViews for SVARs if it is felt that weak instruments are present. It is known that the Anderson-Rubin test is a good way to test hypotheses in the presence of weak instruments and MacKinnon and Davidson (2010) argue that a "wild bootstrap" can produce better outcomes than the Anderson-Rubin statistic. We should also caution that bootstrap methods are not a complete solution here as they are not guaranteed to perform well when instruments are weak. However they would be better than using the asymptotic theory. In general the bootstrap is better than asymptotics if the issue is a "divisor" problem rather than a weak instrument problem, i.e. if it is the distribution of $\frac{\hat{a}_{12}^0}{1-\hat{a}_{12}^0 \hat{a}_{21}^0}$ rather than that of $\hat{a}_{ij}^0$ which causes the problems. There are more general methods to handle weak instruments, but these relate to testing hypotheses about a_{ij}^0 and not to functions like impulse responses.

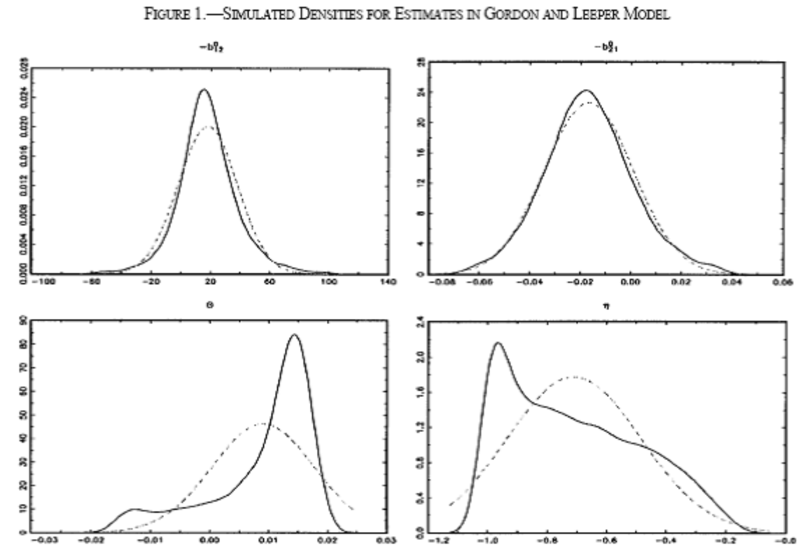


Figure 4.31: Simulated Density Functions from the Parameter Estimators of the Money Demand and Supply Model of Leeper and Gordon

4.8 Other Estimation Methods for SVARs

4.8.1 Bayesian

If the SVAR is exactly identified then the forecasts made with a SVAR would be identical to that from the underlying VAR, i.e. it is only an estimate of B_1 that is important for the forecast and not A_0 and A_1. In this context all that an SVAR provides is an interpretation of the forecast in terms of the shocks indentified with the SVAR. Now this changes if some priors can be placed upon A_0 and A_1so that a BSVAR can be estimated. [19] So we need to consider firstly how one would derive a posterior for these parameters given some priors and, secondly, what sort of priors might be used. Sims and Zha (1998) noted there was a difficulty in that priors such as the Minnesota mentioned in Section 3.4 were generally about B_1. Because this is equal to $A_0^{-1}A_1$ there now needs to be a joint prior about the two matrices. They wrote $p(A_0,A_1) = p(A_1|A_0)p(A_0)$ and prescribed priors for $p(A_1|A_0)$ and $p(A_0)$. In particular $p(A_1|A_0)$ was given the Minnesota form meaning that parameters λ_0,λ_1, and λ_3 need to be prescribed. There is no λ_2 because the SVAR model means that a distinction between the prior variances on own lags versus others is not very meaningful. Because the covariance matrix of the VAR errors e_t is $A_0^{-1}(A_0^{-1})'$ the residual variance prior essentially relates to A_0. In Eviews work this can either be Wishart or an

[19]More generally it will be A_0 and A_1,...,A_p but for simplicity we just use a SVAR(1).

Table 4.1: Forecasting Performance of the the Small Macro Model using Bayesian Estimation Methods, 1998:1-2000:1

Prior ($\mu_5 = \mu_6 = 0$)	Variable	RMSE	MAE
Sims/Zha Normal-Wishart	$Infl$	.803	.777
	Gap	1.162	.973
Sims/Zha Normal-Flat	$infl$	1.272	1.221
	Gap	1.333	1.082

uninformative (i.e., flat) prior. Restrictions on the type of stationarity and co-trending behavior are implemented with the dummy variable priors described in Section 3.4. With these choices of priors one could find posterior densities for $A_1|A_0$ that were normal.

Table 4.1shows the forecasting performance for the small macro mode using the Sims-Zha prior This should be compared to Table 3.1 in Section 3.4.4. There is a slight improvement in the forecasts for inflation over those with the VAR when using the Sims-Zha Normal-Wishart priors.

It is worth noting that there is a tendency today to perform Bayesian estimation in a different way than above, so as to allow more flexibility in the choice of prior. Letting the log likelihood for a SVAR model be $L(\theta)$, where θ are the parameters to be estimated, then the posterior density is the product of the joint density $f(z_{1,..}, z_T|\theta)$ of the data with the prior $p(\theta)$. Hence the log of the posterior density is $C(\theta) = L(\theta) + \log p(\theta)$. Maximizing $C(\theta)$ with respect to θ will give estimates of $\hat{\theta}$ that are the mode of the posterior. Asymptotically the posterior density will be normal so we might assume that the density of $\hat{\theta}$ for any given sample is $N(\theta_{\text{mode}}, (\frac{\partial^2 C}{\partial \theta^2})^{-1})$. In fact this is normally taken to be the proposal density and the actual posterior density of $\hat{\theta}$ is simulated. EViews does not perform Bayesian estimation for arbitrary likelihood functions $L(\theta)$. However, using the *optimize()* routine in EViews applied to $C(\theta)$ would mean that one might obtain the mode and then use the normal density approximation. The addition of $\log p(\theta)$ to the log likelihood will often mean that it will be easier to maximize $C(\theta)$ than to maximize $L(\theta)$, owing to the fact that the prior is a smooth function of θ.

4.8.2 Using Higher Order Moment Information

Only the first two moments have been used to determine the VAR parameters. If the data was normally distributed with constant variance then there would be no further information in the summative model that could be exploited to identify the structure. But, if there is non-normality or changing variances, then the summative model capturing such features provides extra information to identify the parameters. There have been a number of proposals along these lines, e.g. Rigobon (2003) and Lanne and Lutkepohl (2008).

Rigobon used information about breaks in the *unconditional* variance. From the connection between VAR and SVAR shocks we have

$$e_t = A_0^{-1} B\eta_t = \bar{A}\eta_t,$$
$$\eta_t \sim n.i.d(0,\ I_n)$$
$$cov(e_t) = \Omega = \bar{A}\bar{A}',$$

and to this point the last relation has been used to determine $\bar{A} = A_0^{-1}$ after some restrictions (like triangularity) are imposed on A_0.

Now suppose that there is knowledge that a break in the unconditional variance occurs at a time R. This means

(i) for $t = 1, ..., R,\ \eta_t \sim n.i.d(0, I_n), cov(e_t) = \Omega_1$

(ii) for $t = R+1, .., T,\ \eta_t \sim n.i.d(0, D),\ D$ diagonal, $cov(e_t) = \Omega_2$,

resulting in

$$\begin{aligned} \Omega_1 &= \bar{A}\bar{A}' \\ \Omega_2 &= \bar{A}D\bar{A}'. \end{aligned}$$

There are n^2 parameters in $\bar{A}$, and n in D. To estimate these there are $\frac{n(n+1)}{2} + \frac{n(n+1)}{2}$ elements in Σ_1 and Σ_2 (i.e., two covariance matrices) from the summative model. Hence all the elements in $\bar{A}$ can be estimated and no restrictions need to be placed on the structure of $\bar{A}$. This enables one to test the validity of the recursive model since that places restrictions upon $\bar{A}$.

The idea is a clever one, but clearly the timing of the break in $cov(e_t)$ needs to be known, and there must be no shifts in A_0 for it to work. It is not entirely clear why we would see one rather than the other. To implement this estimator, one performs a simultaneous diagonalization on Σ_1 and Σ_2, and this can be done with a generalized singular value decomposition rather than applying the standard one that corresponds to the Cholesky decomposition.

In the example above, there was a break in the unconditional volatility at a known time, i.e. the location of the two regimes is known. But one could also work with a model where data selected the regimes and there was actually no break in the unconditional volatility. An example would be a Markov Switching model with regime dependent volatility. One determines where each regime holds, and then the regime specific variances Ω_1 and Ω_2 are estimated by averaging the squared residuals over the observations for which each regime applies. This idea is used in Herwartz and Lutkepohl (2011). They set up a model where A_0 and the transition probabilities of the MS model are estimated jointly. They report that "The likelihood function is highly non-linear...The objective function has several local optima", and that a very good numerical algorithm was needed to get to the global maximum. This is a feature of many Markov Switching models.

Just as in the breaks-in-variance case, there may be other ways to find the extra equations which will allow the determination of more elements of $\bar{A}$ than a recursive model permits. For example, these extra equations might come from either GARCH structures or non-normality in the errors.

4.8.3 Imposing Independence on the Shocks

One might return to where we started and observe that an alternative to the structural shocks ε_{jt} being uncorrelated is to assume that they are independent of each other. Gourieroux and Monfort (2014) argued that when the SVAR shocks ε_t are linearly related to the VAR shocks i.e. ε_t=He_t, and are independently but not normally distributed, then H will be unique i.e. there is only one linear combination H that will be compatible with the structural shocks being uncorrelated and independent, even though there are many that make them just uncorrelated. Intuitively, this is because independence requires that ε_{it}^k and ε_{jt}^k be uncorrelated for all *k,l=1,...,*∞ (after mean corrections) and so this would rule out many models i.e. many choices of H. For any given H we can compute ε_t and then test for whether the structural shocks are independent, rejecting any H for which this does not hold. This of course requires a test for independence and, in practice, that is not unique. The power of tests for independence can be very weak and the method fails if ε_t normal. So one might ask whether we would expect that data is normally distributed. With real data normality may be acceptable but financial data generally show a lack of independence (as seen with the prevalence of GARCH errors) so the idea of utilizing independence is appealing.

The idea has been built on in different directions. Lanne et al. (2015) also show that the assumption of independent shocks (with at most one marginal distribution being Gaussian) allows a unique recovery of H. Their applied work features some financial data and they assume that ε_t are formed from mixtures of Student t densities. This enables them to estimate H by MLE. Of course there may be other densities that could be used and so just selecting a single one could result in a specification error, but presumably this can be can be tested for. Herwartz and Plodt (2016) assume instead that one would choose the model (H) which produces the least degree of dependence. This is done by setting up some statistical test of independence and then choosing the model that has the greatest p-value for the statistic, since this would imply that the probability of rejecting the null hypothesis of independence is lowest.

These are useful ideas. There is an argument that independence is what is needed if we are to indulge in experiments in which one shock is varied and the others remain constant. Moreover, in cases where financial series such as interest rates and exchange rates are present in SVARs, exploiting the higher order moments may allow us to avoid assumptions such as H being triangular i.e. assuming the SVAR to be recursive. Such data is likely to exhibit non-normality and it is implausible to assume that financial series such as these are contemporaneously unrelated. The biggest issue would seem to be the use of higher order moment information. To model complex densities generally requires large sample sizes and this is rarely the situation in macroeconomics.

Chapter 5

SVARs with I(0) Variables and Sign Restrictions

5.1 Introduction

So far we have discussed methods of estimating the parameters of SVAR systems that impose *parametric* restrictions either directly upon the structure itself or upon the impulse responses. In the past decade a new method has seen increasing use - that of imposing *sign restrictions* upon impulse responses, with early studies being Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005). Table 5.1, taken from Fry and Pagan (2011), gives a partial summary of the studies that have been done with the method. Our aim is to explain the methodology and then evaluate what it can and cannot do. To accomplish the former task Section 5.2 looks at two models that have been used for illustrative purposes in earlier chapters - the market model and the small macro model. Sign restrictions are stated which would be plausible for the types of shocks embedded in the models. Section 5.3 then looks at two methodologies for finding a set of impulse responses that would be compatible with the sign restrictions, and applies these methods to the two simple models used in Section 5.2. Section 5.4 then explores the pros and cons of using sign restrictions, with particular emphasis on some of the difficulties in using sign restrictions to find specific shocks in SVARs. Section 5.5 discusses what happens if there is block exogeneity in the SVAR system and Section 5.6 sets out how standard errors are computed for impulse responses distinguished by sign restrictions.

We will argue that there are four problems that need to be resolved in using the methodology.

1. Sign restrictions solve the *structural identification problem but not the model identification problem.*

2. The lack of a unique model raises questions of which one you choose, and existing methods may not choose a model that is even close to the correct

one. Moreover, many of the schemes for selecting a representative model depend on the way in which the set of models is generated and this may influence the choice.

3. By themselves sign restrictions fail to identify the *magnitude* of the shocks, but it is possible to correct for this by providing a suitable model structure, specifically by providing a normalization of the structural equations.

4. There is a multiple shocks problem that always needs to be addressed.

Although we will refer to the "signs" of IRFs, this is misleading. Any restrictions involving something that can be computed from the structural model, e.g. signs of parameters and covariances, quantitative constraints on these same quantities etc. can all be handled with the same methodology. All that is needed is the ability to be able to compute the quantity numerically.

Table 5.1: Summary of Empirical VAR Studies Employing Sign Restrictions

Fluctuations	Peersman (2005) STNI
	Rüffer *et al.* (2007) STNI
	Sanchez (2007) STNF
Ex Rate	An (2006) STOI
	Farrant/Peersman (2006) STNF
	Lewis (2007) STNF
	Bjørnland/Halvorsen (2008) MTNI
	Scholl/Uhlig (2008) STNI
Fiscal Policy	Mountford/Uhlig (2005, 2008) STNI
	Dungey/Fry (2009) MPTNI
Housing	Jarociński/Smets (2008) MTNI
	Vargas-Silva (2008) STOI
Monetary Policy	Faust (1998) STOI
	Canova/De Nicoló (2002) STOF
	Mountford (2005) STNI
	Uhlig (2005) STOI
	Rafiq/Mallick (2008) STOI
	Scholl/Uhlig (2008) STNI
Technology	Francis/Owyang/Theodorou (2003) MPTOI
	Francis/Owyang/Roush (2005) MPTOF
	Dedola/Neri (2006) SPTOF
	Chari/Kehoe/McGrattan (2008) MPTNF
	Peersman/Straub (2009) STNF
Various	Hau and Rey (2004) STNF
	Eickmeier/Hofmann/Worms (2009) STNI
	Fujita(2009) STOI

Restriction Type: S= Sign only, M = Mixed

Shock Types: P=Permanent, T= Transitory

Number of Shocks: O=One only, N=Numerous

Restriction Source: F=Formal, I= Informal

5.2 The Simple Structural Models Again and Their Sign Restrictions

Again we will use the two simple models to illustrate the arguments. One of these is the market model, which will be written in SVAR(1) form as

$$\begin{aligned} q_t &= -\beta p_t + \phi_{qq} q_{t-1} + \phi_{qp} p_{t-1} + \varepsilon_{D,t} \\ q_t &= \alpha p_t + \phi_{pq} q_{t-1} + \phi_{pp} p_{t-1} + \varepsilon_{St}, \end{aligned}$$

where q_t is quantity, p_t is price, $\varepsilon_{Dt} \sim i.i.d(0, \sigma_D^2)$ is a demand shock, $\varepsilon_{St} \sim i.i.d(0, \sigma_S^2)$ is a supply shock and $cov(\varepsilon_{Dt}, \varepsilon_{St}) = 0$. We think of a supply shock as being a cost shock.

The VAR associated with this model is

$$\begin{aligned} q_t &= b_{qq} q_{t-1} + b_{qp} p_{t-1} + e_{1t} \\ p_t &= b_{pq} q_{t-1} + b_{pp} p_{t-1} + e_{2t}. \end{aligned}$$

Now we would probably expect that the signs of the contemporaneous responses of quantity and prices to *positive* demand and cost shocks would be those of Table 5.2.

Table 5.2: Sign Restrictions for Market Model (Demand/Supply Shocks)

Variable\Shock	Demand	Cost
p_t	+	+
q_t	+	-

One has to be a little careful in applying the restrictions. Take the market model. Then the sign restrictions $\begin{bmatrix} - & - \\ - & + \end{bmatrix}$ would still be interpretable as demand and supply shocks but now they are *negative* rather than positive. So if we came across that pattern these would still show demand and supply shocks. Obviously $\begin{bmatrix} + & - \\ + & + \end{bmatrix}$ and $\begin{bmatrix} - & + \\ - & - \end{bmatrix}$ would also be acceptable. As the number of shocks grows there will be many possible combinations, so it gets rather messy to check all these. For this reason it would seem sensible to keep the number of shocks identified by sign restrictions to a small number.

The small macro model involves an output gap (y_t), inflation (π_t) and a policy interest rate (i_t) and its SVAR form is

$$\begin{aligned} y_t &= x'_{t-1}\gamma_y + \beta_{yi} i_t + \beta_{y\pi}\pi_t + \varepsilon_{yt} \\ \pi_t &= x'_{t-1}\gamma_\pi + \beta_{\pi i} i_t + \beta_{\pi y} y_t + \varepsilon_{\pi t} \\ i_t &= x'_{t-1}\gamma_i + \beta_{iy} y_t + \beta_{i\pi}\pi_t + \varepsilon_{it}, \end{aligned}$$

with a VAR for $x'_t = \begin{pmatrix} y_t & \pi_t & i_t \end{pmatrix}$ of

$$\begin{aligned} y_t &= x'_{t-1}\alpha_y + e_{1t} \\ \pi_t &= x'_{t-1}\alpha_\pi + e_{2t} \\ i_t &= x'_{t-1}\alpha_i + e_{3t}. \end{aligned}$$

In turn we might expect the sign restrictions of Table 5.2 to hold for positive shocks.

Table 5.3: Sign Restrictions for Macro Model Shocks

Variable\Shock	Demand	Cost-Push	Interest Rate
y_t	+	-	-
π_t	+	+	-
i_t	+	+	+

5.3 How Do we Use Sign Restriction Information?

There are two methods for utilizing sign restriction information to find impulse responses to shocks. Both work with a set of uncorrelated shocks. In the first step of both methods impulse responses for uncorrelated shocks ε_t are generated. Then, in the second step, these are judged by whether they have the expected signs of impulses. Those that pass this test are retained. The process is repeated many times, after which there will be many sets of impulse responses that satisfy the signs, and these will generally need to be summarized in some way. Our first method will find many sets of impulse responses by re-combining an initial set of responses, and we will designate this approach as *SRR*, where the *R* stands for re-combination. In the second method the sets of impulse responses are found by varying the A_0 matrix, recognizing that not all of its parameters are estimable from the data. What will vary and produce a large set of impulse responses are the non-estimable coefficients in A_0. The estimable parameters are estimated in such a way as to produce uncorrelated shocks. Because of the emphasis upon the A_0 coefficients of the SVAR we designate this method as *SRC*, where *C* stands for coefficients.

5.3.1 The SRR Method

The key to the SSR method is the selection of a set of *base shocks* η_t that are uncorrelated and which have zero mean and unit variance. One way of getting these is to use the estimated structural shocks from assuming that the system is recursive (this may be totally wrong but all we are trying to do is to get a set of basis shocks that are uncorrelated). In that case

$$A_0^{recur} z_t = A_1 z_{t-1} + \varepsilon_t^R,$$

where A_0^{recur} is a triangular matrix with unity on the diagonals (the equations are normalized) and the ε_t^R are the recursive system structural shocks. Then, the estimated standard deviations of ε_t^R can be used to produce $\bar{\varepsilon}_{jt}^R = \frac{\varepsilon_{jt}^R}{std(\varepsilon_{jt}^R)}$, and the $\bar{\varepsilon}_{jt}^R$ will have unit variances. Consequently if η_t is set equal to $\bar{\varepsilon}_t^R$, it can be thought of as $i.i.d.(0, I_n)$.[1]

Once η_t is found there is an MA structure that determines the impulse responses. Thus, for the recursive model

$$\begin{aligned} z_t &= C^{recur}(L)\varepsilon_t^R \\ &= C^R(L)\bar{\varepsilon}_t^R = C^R(L)\eta_t, \end{aligned}$$

[1] This is not the only way of getting η_t. Suppose we have a VAR(1) in p_t, q_t and the covariance matrix of the errors in the VAR, e_t, is Ω. Applying a singular value decomposition to Ω would produce $P'\Omega P = D$, where D is a diagonal matrix. Consequently $D^{-1/2}P'\Omega D^{-1/2} = I$ and $\eta_t = D^{-1/2}Pe_t$ would have the desired properties. It is easy to find P, D in Matlab and Gauss since $F = D^{-1/2}P'$ is found from the Cholesky Decomposition of Ω. Hence, for any summative model for which one can get e_t, one could apply the Cholesky decomposition to its covariance matrix and thereby create a set of base shocks η_t.

showing that the impulse responses to the shocks η_t are different to the original set. Given this feature the methodology of *SRR* involves forming new shocks by combining those from the base shocks in such a way that the new shocks remain uncorrelated, i.e. $\eta_t^* = Q\eta_t$, where the $n \times n$ matrix Q is required to have the property

$$Q'Q = I_n, QQ' = I_n. \tag{5.1}$$

It is crucial to observe that the new shocks η_t^* need not come from a recursive system even if η_t does. Note that one example of a Q would be the new shocks found by re-ordering the variables in a recursive system.

Why do we have the two restrictions upon Q? The second is used to ensure that the new shocks are also uncorrelated since

$$var(\eta_t^*) = Qvar(\eta_t)Q' = QQ' = I_n.$$

To see the role of the first observe that

$$\begin{aligned} z_t &= C^R(L)\eta_t \\ &= C^R(L)Q'Q\eta_t \\ &= C^*(L)\eta_t^*, \end{aligned}$$

Therefore, after re-combination we have produced a new set of impulse responses C_j^*, but now to shocks η_t^*.

There are a number of ways to find a Q with the required properties. Two popular methods derive from Givens rotations and Householder transforms. We stress now that Q is not unique and this gives rise to what we will term the *model identification issue*. Any given Q produces a new set of shocks and so a new *model*. The models are *observationally equivalent* since $var(z_t)$ is the same. To see that put $B_1 = 0$ and then $var(z_t) = C_0 C_0^{'} = C_0 QQ' C_0 = C_0^* C_0^{*'}$.

The process doesn't end with this first η_t^*. It is iterated to produce many impulse responses by varying Q. Each time these impulses are formed they are tested for whether they obey the maintained sign restrictions. Thus, this leads to the following *modus operandi* for SVARs found from sign restrictions.

1. Start with a set of uncorrelated shocks η_t that have I_n as their covariance matrix.

2. Generate a new set of shocks $\eta_t^* = Q\eta_t$ using a Q with the properties $Q'Q = QQ' = I_n$.

3. Compute the IRF's for this set of shocks.

4. If they have the correct signs RETAIN them. If not, discard the impulse response functions and draw another Q.

5.3.1.1 Finding the Orthogonal Matrices

As mentioned above the matrices Q can be found in a number of ways. A useful choice for expository purposes is that of the Givens matrix.

Givens Matrices A Givens matrix has a particular structure involving cosine and sign terms. When there are two variables ($n = 2$) it has the form

$$Q = \begin{bmatrix} \cos\lambda & -\sin\lambda \\ \sin\lambda & \cos\lambda \end{bmatrix}, 0 \leq \lambda \leq \pi$$

Using $\cos^2\lambda + \sin^2\lambda = 1$ it is easy to see that $Q'Q = I_2$, as required. So, if we put $\lambda = \frac{\pi}{10} = .314$, this gives $Q = \begin{bmatrix} .951 & -.309 \\ .309 & .951 \end{bmatrix}$, and the new shocks η_t^* will be formed from the base ones in the following way:

$$\begin{aligned} \eta_{1t}^* &= .951\eta_{1t} - .309\eta_{2t} \\ \eta_{1t}^* &= .309\eta_{1t} + .951\eta_{2t}. \end{aligned}$$

Consequently, many different Q matrices and impulse responses will be generated by using a range of values for λ. Since λ lies between 0 and π we could just set up a grid of values. An alternative is to use a random number generator drawing λ (say) from a uniform density over 0 to π. Let the *m'th* draw give $\lambda^{(m)}, m = 1, ..., M$. Once a $\lambda^{(m)}$ is available then $Q^{(m)}$ can be computed and there will be M models with IRFs $C_j^{(m)}$. Of course, although all these models are distinguished by different numerical values for λ, they are observationally equivalent, in that they produce an exact fit to the variance of the data on z_t.[2] Only those $Q^{(m)}$ producing shocks that agree with the maintained sign restrictions would be retained.

In the context of a 3 variable VAR (as in the small macro model) a 3×3 Givens matrix Q_{12} has the form

$$Q_{12} = \begin{bmatrix} \cos\lambda & -\sin\lambda & 0 \\ \sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

i.e. the matrix is the identity matrix in which the block consisting of the first and second columns and rows has been replaced by cosine and sine terms and λ lies between 0 and π.[3]

Q_{12} is called a Givens rotation. Then $Q_{12}'Q_{12} = I_3$ using the fact that $\cos^2\lambda + \sin^2\lambda = 1$. There are three possible Givens rotations for a three variable system - the others being Q_{13} and Q_{23}. Each of the Q_{ij} depends on a *separate* parameter λ_k $(k = 1, .., 3)$. In practice most users of the approach have adopted the multiple of the basic set of Givens matrices as Q. For example, in the three variable case we would use

$$Q_G(\lambda) = Q_{12}(\lambda_1) \times Q_{13}(\lambda_2) \times Q_{23}(\lambda_3).$$

[2] This statement assumes a zero mean for z_t.

[3] In general Q_{ij} is formed by taking an $n \times n$ identity matrix and setting $Q_{ij}^{ii} = \cos\lambda$, $Q_{ij}^{ij} = -\sin\lambda$, $Q_{ij}^{ji} = \sin\lambda, Q_{ij}^{jj} = \cos\lambda$, where the superscripts refer to the row and column of Q_{ij}.

It's clear that Q_G is orthogonal and so shocks formed as $\eta_t^* = Q_G\eta_t$ will be uncorrelated. Because the matrix Q_G above depends upon three different λ_k one could draw each λ_k from a $U(0,\pi)$ density function.

Householder Transformations A suitable Q matrix can also be found using Householder transforms. In the three variable case one generates a 3×3 matrix W from a three dimensional multivariate normal with zero mean and covariance matrix I_3. Then the QR decomposition is applied to W. The QR decomposition is available in MATLAB, GAUSS, Stata and decomposes W as $W = Q_R R$, where Q_R is unitary and R is triangular, so that Q_R can be used as a Q. The method is computationally efficient relative to Givens for large n and is numerically easy to perform. It was first proposed by Rubio-Ramírez *et al.* (2006).

5.3.2 The SRC method

Because the shocks are connected to a SVAR, and we know from Chapter 4 that only $\frac{n(n-1)}{2}$ elements of A_0 can be estimated (when A_0 has unity on the diagonals), after using the restriction that the shocks are uncorrelated there will be $\frac{n(n-1)}{2}$ non-estimable parameters in A_0. These need to be fixed to some values if estimation is to proceed. The idea behind the *SRC* approach is to choose some values for the non-estimable parameters in A_0 and to then estimate the remainder with a method which ensures that the shocks are uncorrelated. Because there is no unique way to set values for the non-estimable parameters, there will be many impulse responses coming from changing these parameter values, i.e. it performs the same task as varying the Q values in SRR. So the key to the methodology resides in generating many values for the non-estimable parameters, and these will be taken to depend upon some quantities designated as θ. Broadly we will find values for the non-estimable parameters by generating candidate values for θ from a random number generator. The context may determine exactly how that would be done. Once again the models found with different values of θ are observationally equivalent as the SVAR is exactly identified.

5.3.3 The SRC and SRR Methods Applied to a Market Model

Because the sign restrictions relate to contemporaneous responses it is useful to omit any dynamics from the simple market model. Thus it will have the form

$$q_t = \alpha p_t + \varepsilon_{1t} \tag{5.2}$$

$$q_t = -\beta p_t + \varepsilon_{2t}, \tag{5.3}$$

where q_t is quantity, p_t is price, and the shocks ε_{jt} are $n.i.d.(0,\sigma_j^2)$ and uncorrelated with one another. The first curve might be a supply curve and the second a demand curve (implying that both α and β are positive). Because lags are omitted from (5.2) and (5.3) this is a structural system but not an SVAR. Nevertheless, it is useful to abstract from lags, and this can be done without loss of generality. Based on this model we could form

$$\begin{aligned} \sigma_S^{-1} q_t &= (\alpha/\sigma_S) p_t + \eta_{1t} & (5.4) \\ \sigma_D^{-1} q_t &= -(\beta/\sigma_D) p_t + \eta_{2t}, & (5.5) \end{aligned}$$

where η_{jt} are $n.i.d(0,1)$. The model can be represented as

$$\begin{aligned} a q_t &= b p_t + \eta_{1t} & (5.6) \\ c q_t &= d p_t + \eta_{2t}. & (5.7) \end{aligned}$$

A unit shock to the ε_{jt} is then equivalent to one standard deviation shocks in supply ($e_{S,t} = \sigma_S \eta_{1t} = \varepsilon_{1t}$) and demand ($e_{D,t} = \sigma_D \eta_{2t} = \varepsilon_{2t}$). The corresponding impulse responses to these shocks will be $\begin{bmatrix} a & -b \\ c & -d \end{bmatrix}^{-1}$.

5.3.3.1 The SRR Method Applied to the Market Model

In the standard sign restrictions methodology (SRR) one way to initiate the process is to start with a recursive model. For the market model this could be

$$\begin{aligned} q_t &= s_1 \eta_{1t} & (5.8) \\ p_t &= \phi q_t + s_2 \eta_{2t}. & (5.9) \end{aligned}$$

Data is that on q_t and p_t and the η_{jt} are $n.i.d(0,1)$, with s_j being the standard deviations of the errors for the two equations. The first stage of SRR is then implemented by applying some weighting matrix Q to the initial shocks η_{1t} and η_{2t} so as to produce new shocks η_{1t}^* and η_{2t}^*, i.e. $\eta_t^* = Q\eta_t$. As mentioned above Q is chosen in such a way as to ensure that $QQ' = Q'Q = I$, which means that the new shocks are also uncorrelated with unit variances. One matrix to do this is the Givens matrix $Q = \begin{bmatrix} \cos\lambda & -\sin\lambda \\ \sin\lambda & \cos\lambda \end{bmatrix}$, where λ are values drawn from the range $(0, \pi)$. After adopting this the new shocks $\eta_t^* = Q\eta_t$ will be

$$\begin{aligned} \cos\lambda\eta_{1t} - \sin\lambda\eta_{2t} &= \eta_{1t}^* \\ \sin\lambda\eta_{1t} + \cos\lambda\eta_{2t} &= \eta_{2t}^*. \end{aligned}$$

Using the expressions for η_{1t} and η_{2t} in (5.8) - (5.9) we would have

$$\begin{aligned} (\cos(\lambda)/s_1)q_t - (\sin(\lambda)/s_2)(p_t - \phi q_t) &= \eta_{1t}^* \\ (\sin(\lambda)/s_1)q_t + (\cos(\lambda)/s_2)(p_t - \phi q_t) &= \eta_{2t}^*, \end{aligned}$$

which, after re-arrangement, is

$$\begin{aligned} [(\cos(\lambda)/s_1) + \sin(\lambda)(\phi/s_2)]q_t - (\sin(\lambda)/s_2)p_t &= \eta_{1t}^* & (5.10) \\ [(\sin(\lambda)/s_1) - \cos(\lambda)(\phi/s_2)]q_t + (\cos(\lambda)/s_2)p_t &= \eta_{2t}^*. & (5.11) \end{aligned}$$

Now this has the same form as (5.6) - (5.7) when

$$\begin{array}{rr} a = (\cos(\lambda)/s_1) + \sin(\lambda)(\phi/s_2) & b = (\sin(\lambda)/s_2) \\ c = (\sin(\lambda)/s_1) - \cos(\lambda)(\theta/s_2) & d = -(\cos(\lambda)/s_2) \\ \varepsilon_{jt} = \eta^*_{jt} & \end{array} \tag{5.12}$$

The latter can hold since both sets of random variables are uncorrelated and *n.i.d.*

Now the impulse responses for η^*_t are produced by re-combining those for η_t with the matrix Q, and this is generally how the strategy employed in SRR is described. An alternative view would be that the SRR method generates many impulse responses by expressing the A_0 coefficients of the SVAR model in terms of λ, and then varying λ over the region $(0, \pi)$. Once the impulse responses are found sign restrictions are applied to decide which are to be retained. So we are generating many impulse responses by making the market model parameters A_0 depend upon λ and the data (through ϕ, s_1 and s_2).

5.3.3.2 The SRC Method Applied to the Market Model

Rather than expressing the model parameters in terms of λ, consider the possibility of going back to (5.2) and making the coefficient α (the non-estimable one) a function of θ, where θ varies over a suitable range. Given a value for θ this will fix α. The estimable coefficients then need to be found from the data in such a way as to produce uncorrelated shocks.

After setting θ to some value θ^* estimation can be done using the following method.

1. Form residuals $\hat{\varepsilon}^*_{1t} = q_t - \alpha(\theta^*)p_t$.
2. Estimate σ_1 with $\hat{\sigma}^*_1$, the standard deviation of these residuals.
3. Using $\hat{\varepsilon}^*_{1t}$ as an instrument for p_t estimate β by Instrumental Variables (IV) to get $\hat{\beta}^*$.
4. Using $\hat{\beta}^*$ form the residuals $\hat{\varepsilon}^*_{2t} = q_t + \hat{\beta}^* p_t$. The standard deviation of these, $\hat{\sigma}^*_2$, will estimate the standard deviation of the second shock. By the nature of the estimation procedure the shocks $\hat{\varepsilon}^*_{1t}$ and $\hat{\varepsilon}^*_{2t}$ are orthogonal.

Using earlier results, the contemporaneous impulse responses to one standard deviation shocks will be $\begin{bmatrix} 1 & -\alpha(\theta^*) \\ 1 & \hat{\beta}^* \end{bmatrix}^{-1} \begin{bmatrix} \hat{\sigma}^*_1 & 0 \\ 0 & \hat{\sigma}^*_2 \end{bmatrix}$. Accordingly, just as happened with λ in the SRR approach, we can vary θ and thereby generate many impulse responses. These are directly comparable with the impulse responses generated by SRR, except that they all depend upon θ and the data (via the IV estimation) rather than λ and the data. Because the technique consists of finding a range of impulse responses by varying the coefficient α (through varying θ) it is the SRC method mentioned earlier. Of course the IV method here just provides a simple explanation of how SRC works. Once $\alpha(\theta^*)$ is formed one could just apply MLE to the system since IV and MLE are identical in this exactly identified system. Sometimes there can be convergence issues with MLE and a number of different starting values are needed, and, if this happens, the IV estimates should be used as the starting values since the MLE must equal them.

5.3.3.3 Comparing the SRR and SRC Methodologies

It is worth looking closer at these two methods. A number of points emerge.

(i) θ will normally be chosen so as to get a range of variation in α that is (-∞, ∞). This can be done by drawing θ from a uniform (-1,1) density and then setting$\alpha = \frac{\theta}{1-abs(\theta)}$. [4]By comparison in SRR λ is drawn from a uniform density over $(0, \pi)$, because of the presence of λ in the harmonic terms. In both approaches one has to decide upon the number of trial values of θ and λ to use, i.e. how many sets of impulse responses are to be computed. We note that there may be cases where it is possible to bound the values of the non-estimable parameters, e.g. restrict θ so that it is less than (say) fifty, and this would then have implications for how θ is generated (or possibly one would simply discard models for which the non-estimable parameters lay outside the bounds).

(ii) In a SVAR with n variables and no parametric restrictions the number of λ_j to be generated in the SRR method equals $n(n-1)/2$. Thus, for $n = 3$, three $\lambda_j' s$ are needed. This is also true of the number of θ_j used in SRC. So problems arising from the dimensions of the system are the same for both methods. It should be noted however that, when parametric restrictions are also applied along with sign restrictions, the number of θ_j may be much smaller, and this will be shown later. Such an outcome should be apparent because parametric restrictions increase the number of estimable parameters in A_0 and, since θ_j relates to the non-estimable parameters, a smaller number of θ_j need to be generated.

5.3.4 Comparing SRC and SRR With a Simulated Market Model

To look more closely at these two methods we simulate data from the following market model[5]

$$q_t = 3p_t + \sqrt{2}\varepsilon_{2t} \qquad (5.13)$$
$$q_t = -p_t + \varepsilon_{1t}$$

The true impulse responses for price and quantity (with the demand shock first and supply second) are $\begin{bmatrix} .75 & .3536 \\ .25 & -.3536 \end{bmatrix}$. Five hundred values for θ and λ were generated from a uniform random number generator (over (0,π) for λ and

[4]In Ouliaris and Pagan (2016) other ways of generating θ are considered. It seemed that this procedure produced the best coverage of the parameter space for α

[5]Notice that we have made ϵ_{1t} the demand equation shock compared to ϵ_{2t} in (5.3).

(-1,1) for θ) and the impulse responses based on the θ and λ were compared to the sign restrictions in Table 5.2. SRR generates impulses that are compatible with the sign restrictions 87.8% of the time and for SRC it is 85.4%. This is a high percentage but, since the model is correct, this is what would be expected. Inspecting these we find that among the 500 impulse responses the closest fit to the true impulse responses for each method was[6]

$$SRC = \begin{bmatrix} .7369 & .3427 \\ .2484 & -.3605 \end{bmatrix}, SRR = \begin{bmatrix} .7648 & .3529 \\ .2472 & -.3563 \end{bmatrix}.$$

It is clear that among the 500 sets of responses for each method there is at least one that gives a good match to the true impulse responses. Changing the parameter values for the market model did not change this conclusion.

There was however some sensitivity to sample size. In the simulation above 1000 observations were used. When it is reduced to 100 observations the equivalent results are

$$SRC = \begin{bmatrix} .7421 & .3615 \\ .1702 & -.3923 \end{bmatrix}, SRR = \begin{bmatrix} .7828 & .4119 \\ .1780 & .-.3806 \end{bmatrix}.$$

It seemed that SRC tended to produce a slightly better fit to the true impulse responses, although they both provide a reasonable match.

5.3.5 Comparing SRC and SRR With a Small Macro Model and Transitory Shocks

We will now look at the two methods in the context of the three variable small macro model.[7] This was also used in Fry and Pagan (2011). The variables in the system consist of three variables y_{1t},y_{2t} and y_{3t}, where y_{1t} is an output gap, y_{2t} is quarterly inflation, and y_{3t} is a nominal interest rate. All variables are assumed to be $I(0)$ and there are three transitory shocks - labeled demand, costs and an interest rate. The expected signs of the contemporaneous impulse responses are given in Table 5.2.

The model fitted is the SVAR(1)[8]

$$\begin{aligned} y_{1t} &= a_{12}^0 y_{2t} + a_{13}^0 y_{3t} + a_{12}^1 y_{2t-1} + a_{13}^1 y_{3t-1} + a_{11}^1 y_{1t-1} + \varepsilon_{1t} && (5.14) \\ y_{2t} &= a_{21}^0 y_{1t} + a_{23}^0 y_{3t} + a_{22}^1 y_{2t-1} + a_{23}^1 y_{3t-1} + a_{21}^1 y_{1t-1} + \varepsilon_{2t} && (5.15) \\ y_{3t} &= a_{31}^0 y_{1t} + a_{32}^0 y_{2t} + a_{32}^1 y_{2t-1} + a_{33}^1 y_{3t-1} + a_{31}^1 y_{1t-1} + \varepsilon_{2t}. && (5.16) \end{aligned}$$

The *SRR* method begins by setting $a_{12}^0 = 0, a_{13}^0 = 0$ and $a_{23}^0 = 0$ to produce a recursive model, and then recombines the impulse responses found from this model using the Q_G matrix that depends upon λ_1, λ_2 and λ_3. In contrast,

[6]We just use a simple Euclidean norm to define the closest match to the true values. The impulse responses are to a one standard deviation shock.

[7]The code for replicating the results in this section can be found in the folder "SIGN" in the files called *src.prg* and *srr.prg.*

[8]For illustration we assume a SVAR of order one, but in the empirical work it is of order two.

the *SRC* method proceeds by first fixing a_{12}^0 and a_{13}^0 to some values and then computing residuals $\hat{\varepsilon}_{1t}$. After this (5.15) is estimated by fixing a_{23}^0 to some value and using $\hat{\varepsilon}_{1t}$ as an instrument for y_{1t}. Lastly the residuals from both (5.14) and (5.15), $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$, are used as instruments for y_{1t} and y_{2t} when estimating (5.16).[9]

Once all shocks have been found impulse responses can be computed. Of course three parameters have been treated as non-estimable and so they need to be generated. This is done by defining $a_{12}^0 = \frac{\theta_1}{(1-abs(\theta_1))}, a_{13}^0 = \frac{\theta_2}{(1-abs(\theta_2))}, a_{23}^0 = \frac{\theta_3}{(1-abs(\theta_3))}$, and then getting realizations of θ_1, θ_2 and θ_3 from a uniform random generator. Note that three different random variables θ_j are needed and these correspond to the three λ_j in the Givens matrices. As for the market model the methods are computationally equivalent.

Unlike the market model it is not easy to find impulse responses that satisfy the sign restrictions. For both methods only around 5% of the impulse responses are retained. 1000 of these were plotted for *SRR* in Figure 1 of Fry and Pagan. Therefore, Figure 5.1 below gives the same number of impulse responses from the *SRC* method (here the positive cost shocks mean a negative productivity shock and, because, Fry and Pagan used a positive productivity shock in their figure, an allowance needs to be made for that when effecting a comparison). It seems as if *SRC* produces a broader range of impulse responses than *SRR*, e.g. the maximal contemporaneous effect of demand on output with SRC is more than twice what it is for SRR (we note that all impulse responses in the ranges for both SRC and SRR are valid in that they have the correct signs and they are all observationally equivalent).[10]

It is clear that there is a large spread of values, i.e. many impulse responses can be found that preserve the sign information and which fit the data equally. The spread here is *across models* and has nothing to do with the variation in data. Hence it is invalid to refer to this range as a "confidence interval" as is often done in the literature. Of course in practice we don't know A_1, Ω and so these need to be estimated, and that will make for a confidence interval. We return to that issue in a later section. Such dependence on the data provides some possible extra variation in the spread for impulse responses, but it doesn't help to conflate this with the variation in them across observationally equivalent models.

It is worth observing that both *SRC* and *SRR* have a potential problem in

[9] Of course since the SVAR is exactly identified this IV procedure is just FIML. The reason for explaining it in terms of IV is that such an approach will be clearer when we come to permanent shocks. Nevertheless, given that programs like EViews and Stata estimate the SVARs by FIML it will generally be easier to just set a_{ij} to generated values and then perform FIML.

[10] This points to the fact that the impulses found with SRC and SRR may not span the same space. Thinking of this in the context of the market model it is clear that we could find an α (for SRC) that would exactly reproduce the same α as coming from SRR. But the estimate of β found by both methods would then differ, and that would lead to different impulse responses. These two sets of impulse responses will be connected by a non-singular transformation but it will vary from trial to trial. If it did not vary then the impulse responses would span the same space.

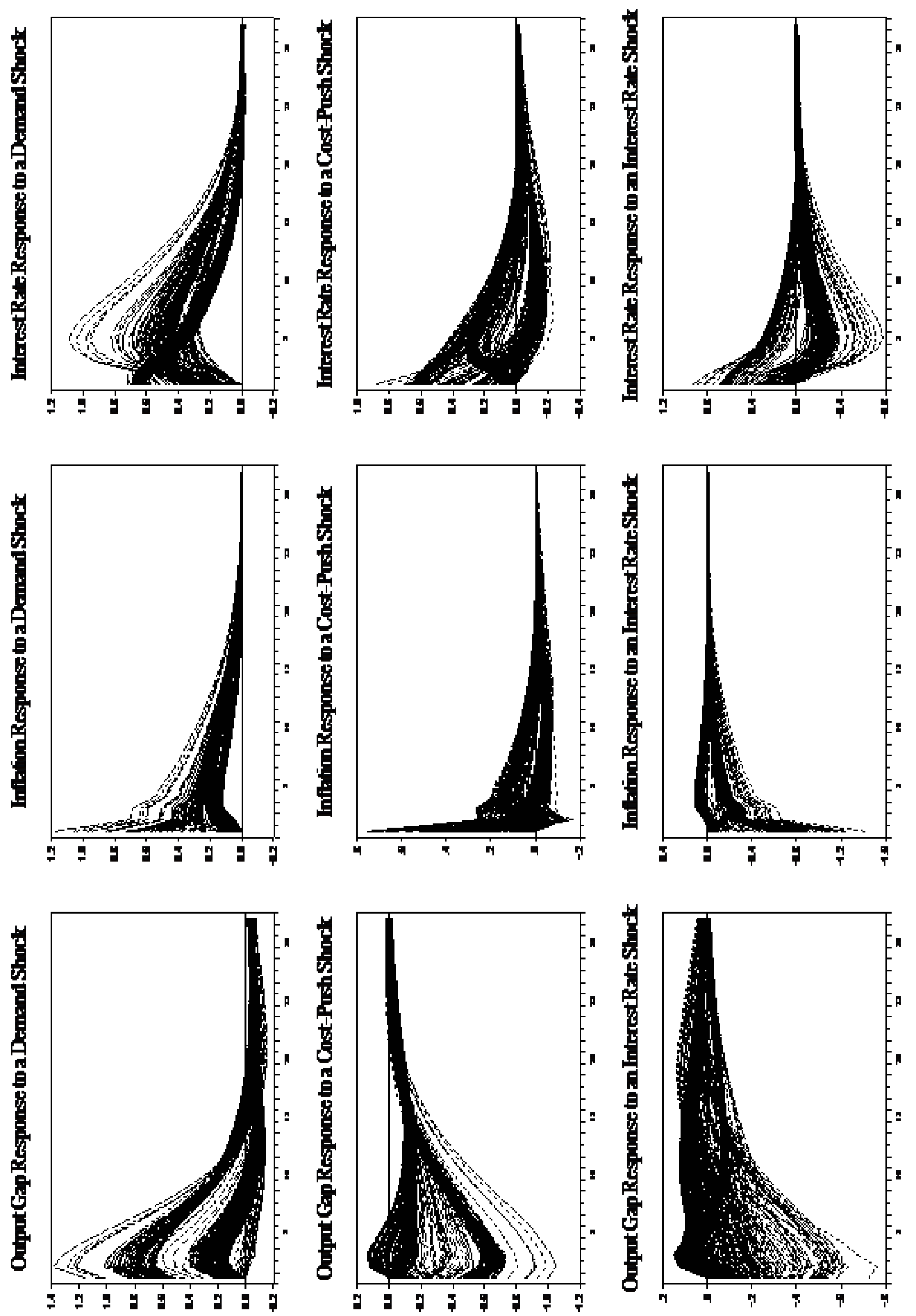


Figure 5.1: 1000 Impulses Responses from SRC Satisfying the Sign Restrictions for the Small Macro Model using the Cho-Moreno Data

generating the widest possible range of impulse responses. For *SRR* this arises in two ways. Firstly, in the selection of the initial set of impulse responses. As mentioned earlier, this has been done by either the Cholesky or singular value decompositions. The Cholesky decomposition requires an ordering of the variables, so there will be different initial impulse responses depending on which ordering one uses. Of course the SVD just adds another set. For any given Q then we would get a different set of impulse responses depending on which choice of factorization is used to initiate the process. Secondly, there is Q itself. The Givens and simulation based method provide a Q with the requisite properties, but there may well be others. If so, then one might expect different impulse responses when those Q matrices are applied to the same initial model. This problem shows up with *SRC* as well. Now it is in terms of the parameters that are taken to be unidentified and which need to be generated. To be more concrete, consider the fact that in our example with the small macro model a^0_{12},a^0_{13} and a^0_{23} were the generated parameters. Instead, one might have chosen a^0_{31},a^0_{32} and a^0_{21}. If so, estimation would have started with (5.16) rather than (5.14). For both methods this is a potential problem but perhaps not a real one (provided the number of trials is large). It may well be that the range of impulse responses generated is much the same, regardless of either the initial choice of impulse responses or the unidentified parameters. What might happen is that some choices require more trials than others in order to produce a relatively complete set of impulse responses. Fundamentally, the issue arises because both SRR and SRC focus on first producing a set of impulse responses to uncorrelated shocks, after which they can be checked to see if they satisfy sign restriction information. However, neither guarantees that this set is exhaustive.

5.4 What Can and Can't Sign Restrictions Do for You?

5.4.1 Sign Restrictions Will Not Give You a Single Model - The Multiple Models Problem

How do we deal with the fact that there are many models that satisfy the sign restrictions? If there is a narrow spread we would be presumably happy to choose a single one. In practice people mostly report the median and some "percentiles" like 5% and 95%. One of the problems with the median can be understood in a model with two shocks and where we look at the first variable. Choose the medians of the impulse responses of that variable to the two shocks and designate them by $C^{(k_1)}_{11} = med\{C^{(k)}_{11}\}$ and $C^{(k_2)}_{12} = med\{C^{(k)}_{12}\}$, where k_1 is the model that has the median of the C_{11} impulses and k_2 is the median for those of C_{12}. In general k_1 doesn't equal k_2 so these median values are impulse responses *from different models.* It is hard to make sense of that. It is like using an impulse response for a money shock from a monetary model and one for a technology shock from an RBC model. Moreover if they come from different models they are no longer uncorrelated as that requires that they be constructed with a common Q. The whole point of SVAR work is to ensure that shocks remain uncorrelated. In the event that they are correlated techniques such as variance decompositions cannot be applied.

Another problem with the median comes from the fact that the set of impulse responses being summarized depends upon how they were generated i.e. how θ and λ are chosen. In its simplest form the problem can be seen in application of SRC to the market model. Here the parameter estimate $\widehat{\beta}$ depends on $\alpha(\theta)$, and so the density of $\widehat{\beta}$ (across models) must depend on the density of θ. Consequently, the median value of impulse responses which come from combining α and $\widehat{\beta}$ vary according to the density chosen for θ. So one needs to recognize that, whilst there is a single median for a given set of impulse responses, this depends on how θ and λ are chosen. This issue has been pointed out by Baumeister and Hamilton (2014) in their critique of Bayesian methods for summarizing the range of impulse responses. What one makes of this depends a good deal on whether one wants to summarize the generated impulse responses with a single measure or whether one is simply interested in the range of outcomes, since that will not be affected by the choice of method for generating θ and λ(although one does need to simulate many of these for that to be true). If we designate the maximum and minimum values of $\widehat{\beta}$ by $\widehat{\beta_{max}}$ and $\widehat{\beta}_{min}$ then it seems that a reasonable way to sumamrize the outcomes in a single number would be to use the average of $\widehat{\beta_{max}}$ and $\widehat{\beta}_{min}$. Provided we have generated many models this should be less sensitive to how $\alpha$is generated than the median or other percentiles would be. Of course the average may not be associated with a single model so we would need to use something like the MT method to choose a model that is as close to it as possible.

If the median is thought to be a desirable benchmark Fry and Pagan (2007) proposed the MT (median target) method which chooses a single model whose impulses were as close as possible to the median values. It is worth computing this since, if it differs a lot from the median impulses, then you know that the impulse responses that are presented must be associated with correlated shocks. Sometimes one sees empirical work where the MT method and medians are close. An example comes from using sign restrictions for demand and supply functions with the Blanchard and Quah (1989) data. For the small macro model however they can be quite different - see Figure 5.2.

Some other methods have been proposed to narrow the range.

1. Uhlig (2005) proposed a criterion that expressed a preference for the largest (in absolute terms) impulses. In this connection he says " it might therefore be desirable to pick the one, which generates a **more decisive response** of the variables, for which sign restrictions are imposed: this is what the penalty-function approach does." (p 414). It is unclear why it is a good idea to select an extreme value from the range.

2. Some investigators use sign restrictions on more than just contemporaneous impulse responses. This can also narrow the range of models

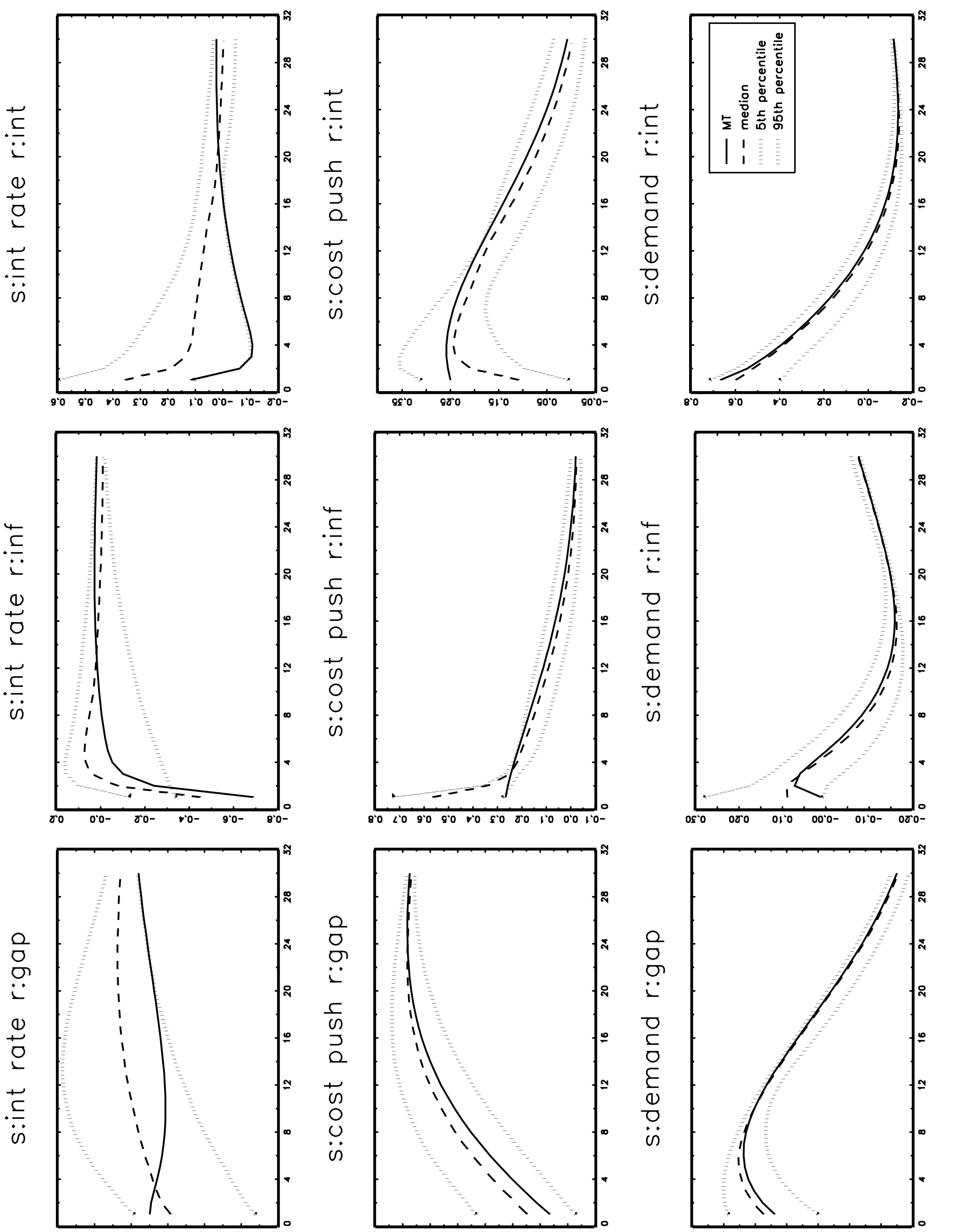


Figure 5.2: Impulse Responses for Sign Restrictions in the Small Macro Model - Median, 5 and 95 percentiles and the MT method

quite sharply, as seen in the 1000 impulse responses from the small macro model. In many cases, however, this approach may not add much because $C_j = D_j C_0$ and, since D_j is fixed from the VAR and does not depend on the structure, whenever $D_0 > 0$, any $C_0 > 0$ automatically ensures that $C_j > 0$.

3. One might reject many of the generated impulse responses as implausible using some additional criteria. Kilian and Murphy (2012, p 1168) refer to these as "...additional economically motivated inequality restrictions" - in their case the magnitude of the estimated oil supply elasticity on impact.

This problem of model identification is not unique to sign restrictions. In recursive systems there are many possible "orderings" of variables and so many possible models, all of which fit the data equally well. Mostly this is dealt with via a comment like "we tried other orderings with similar results", which would seem to imply a narrow range of responses. To make this point more concrete take a recursive market model. Then it needs to be decided whether we would order p_t or q_t first. When you think about this you can see that the ordering is really a statement about how the market operates. In one case quantity is predetermined and price adjusts. In the other, price is predetermined and quantity adjusts. It may be that there is institutional information about the relative likelihood of each of these, i.e. there is *extra information* other than sign restrictions. This example suggests that what is needed in any sign restriction application is supplementary information, perhaps of the sort that Kilian and Murphy use. It is therefore relevant to observe that, if we insisted on independence of as well as a lack of correlation between the shocks, it might be possible to come up with a single model and so to avoid the issue of how one is to summarize the range of impulse responses altogether.

5.4.2 Sign Restrictions and the Size of Shocks?

The SRR process always starts with unit variance shocks that are uncorrelated. Suppose one started with $v_{it} = \frac{\varepsilon_{it}}{\sigma_i}$, where ε_{it} are the true shocks and σ_i are the true standard deviations. Then the base shocks would be $\eta_{it} = v_{it}$. If these gave impulses satisfying the sign restrictions a rise of one unit in η_{it} would mean a rise in ε_{it} of σ_i, i.e. the impulse responses identified by sign restrictions are *for one standard deviation changes* in the true shocks. The problem then is that we don't know what σ_i is unless $\sigma_i = 1$.[11]

In terms of the market model the problem is that $\varepsilon_{Dt} \sim i.i.d(0, \sigma_D^2)$, $\varepsilon_{St} \sim i.i.d(0, \sigma_S^2)$ and, by setting $\eta_{1t} = \sigma_S^{-1}\varepsilon_{St}, \eta_{2t} = \sigma_D^{-1}\varepsilon_{Dt}$, the demand and supply equations have been converted to a structural system that has shocks with a unit variance. What we really want are impulse responses to the demand and

[11] In a recursive system a one standard deviation to a shock can be found by looking at the response of the variable that is the dependent variable of the equation the shock is attached to. But this is not true in non-recursive systems and sign restrictions generate many non-recursive systems. Only if the correct model is recursive would we be able to infer the standard deviation from the response of a model variable.

supply shocks and not to the η_{it}. The latter have the same signs as those for $\varepsilon_{Dt}, \varepsilon_{St}$ but, because sign information is invariant to the "magnitude" of shocks, one does not directly recover the standard deviation of the shocks of interest. Much of the literature seems to treat the impulse responses as if they were responses to a one unit shock, and this is clearly incorrect. How then is it that the standard deviations can be estimated either when parametric restrictions or the SRC method are applied? The answer lies in the normalization used in those methods. Once this is provided the implied structural equations in the SRR method can be recovered, along with the standard deviations of their shocks. To illustrate, take the market model used to simulate data and write it in the form

$$
\begin{aligned}
q_t &= -p_t + \eta_{1t} \\
p_t &= \frac{1}{3}q_t - \frac{\sqrt{2}}{3}\eta_{2t} = \frac{1}{3}q_t - .4714\eta_{2t}.
\end{aligned} \tag{5.17}
$$

Taking the C_0 from SRR that was closest to the true value of the impulse responses for the market model, namely $C_0 = \begin{bmatrix} .7648 & .3529 \\ .2472 & -.3563 \end{bmatrix}$, gives

$C_0^{-1} = \begin{bmatrix} .9905 & .9810 \\ .6872 & -2.1262 \end{bmatrix}$. Thereafter, utilizing $A_0 = C_0^{-1}$, and imposing a normalization one gets the implied relations of

$$
\begin{aligned}
q_t &= -(\frac{.9905}{.9810})p_t + \frac{1}{.9905}\eta_{1t} \\
p_t &= (\frac{.6876}{2.1262})q_t - \frac{1}{2.1262}\varepsilon_{2t} = .32q_t - .4703\eta_{2t}.
\end{aligned} \tag{5.18}
$$

From these equations the standard deviations of the shocks would be 1.01 and .4703, compared to the true ones of 1 and .4714. Of course many impulse responses are produced by SRR and so there will be many values for σ_i. It is worth emphasizing that this is also true for SRC, since that method estimates σ_i as part of the estimable parameter set. Just as impulse responses need to be summarized in some way, this will equally be true of the $\sigma_i^{(m)}$ found for the $m'th$ model. There is not just one single standard deviation unless a particular value for m is chosen by some criterion.

Does the problem just outlined matter? In some cases the answer is in the negative. The *shape* of the true impulse responses does not depend on σ_i. There are also exercises that do not require the standard deviation of the shocks since $z_t = C(L)\varepsilon_t = C(L)\sigma\sigma^{-1}\varepsilon_t = C^*(L)\eta_t^*$, e.g. forecast variance and variable decompositions. Nevertheless, in many cases policy questions do depend on knowing the standard deviation of the shocks, e.g. to answer questions like what is the effect of a 100 basis point shock in interest rates? Another problem arising from not knowing the standard deviation of the true shocks is when a comparison is made (say) of fiscal policy impulses across different countries (or time), since any differences in magnitudes of the impulse responses may be simply due to different standard deviations in the shocks for the countries (time).

5.4.3 Where Do the True Impulse Responses Lie in the Range of Generated Models?

Often sign restrictions are found from some DSGE model. The methodology to do this is to compute impulses from that model for a big range of the model parameters and then use the impulse response functions whose signs are robust to the parameter values. Turning this around, one might ask whether you would recover the true impulse responses if data were simulated from the DSGE model and then the impulse responses were found with a sign restricted SVAR? In general all one can say is that the true impulse responses will lie in the M models that are generated from SRR or SRC (provided of course that M is large enough). But where in the range do they lie? There is nothing which says that they will lie at the median, as that is not "most probable" in any sense. It is just a description of the range of generated impulses. In Fry and Pagan (2011) a macro model was simulated and the true impulses were found to lie at percentiles like 12.5 and .4, not at the median (50^{th}percentile). So there is no reason to think that the median has much to recommend it, except as a description of the range of outcomes. This result was also found in Jääskelä and Jennings (2011).

Doing this for the market model produces median responses of

$$SRC = \begin{bmatrix} .4082 & .3119 \\ .5284 & -.4432 \end{bmatrix}, SRR = \begin{bmatrix} .6234 & .5665 \\ .3429 & -.2655 \end{bmatrix}.$$

Neither coincides with the true values nor are they the same, which can just be a result of λ and θ being generated differently. Indeed, while the median response of price to a demand shock is .4082 for SRC, the true response of .75 lies at the 89th percentile, and so the median response is only around one half of the true value. Unless one had some extra information for preferring one set of impulse responses to another the median has no more appeal than any other percentile. As the result above shows, the percentile at which the true impulse responses lie can also vary with which method, SRC or SRR, is used.

5.4.4 What Do We Do About Multiple Shocks?

There is also a *multiple shocks* problem. Often researchers only want to identify one shock. This means that there will be n uncorrelated shocks but $n-1$ of these are "un-named". As such we know nothing about their impacts. Can one do this? As an illustration of the problems take the market model with two shocks, where it is assumed that the only information known about their impacts is summarized in C_0 as $\begin{bmatrix} + & ? \\ + & ? \end{bmatrix}$, where ? means that no sign information is provided. Clearly one doesn't have enough information here to discriminate between the shocks, as one would not know what to do if the pattern in the generated responses was found to be $\begin{bmatrix} + & + \\ + & + \end{bmatrix}$, as two demand shocks in the same model would be implausible. This problem is sometimes mentioned in applied work but details supplied of what was done about it are often scanty. One suspects that the search terminates when one set of correct sign restrictions is found. Instead the presence of two shocks with the requisite sign restrictions should cause the model to be rejected as it cannot have two shocks with the same sign patterns.

5.4.5 What Can Sign Restrictions do for you?

There seem to be four ways for sign restrictions to be useful.

1. They tell you about the range of possible models (impulse responses) that are compatible with the data.

2. The number of rejections of the generated models because of a failure to match the sign restrictions would seem to be informative. In Peersman's (2005) SVAR model estimated by sign restrictions more than 99.5% of generated models were rejected. This has to make one think that the data is largely incompatible with the sign restrictions.

3. They are good for telling you about the *shapes* of responses and this can help in choosing a parametric model.

4. Sometimes we might be happy to apply parametric restrictions so as to isolate certain shocks but are doubtful about them for isolating others. Because sign restrictions utilize much weaker information, in such cases it can be useful to employ the sign restrictions to capture the remaining shocks. Thus a combination of parametric and sign restrictions could be desirable. As will be shown in the next chapter the SRC method is well designed to handle combinations of parametric and sign restrictions and the SRR approach has also been extended in this way - see Arias *et al.* (2014).

5.5 Sign Restrictions in Systems with Block Exogeneity

Suppose we are using a VAR with exogenous variables, e.g. in an open economy context. Then there are two sets of variables z_{1t} and z_{2t} . For convenience we will refer to the former as foreign variables and the latter as domestic. The original system is recursive but, importantly, no lags of z_{2t} (domestic variables) appear in the VAR equations for z_{1t} (foreign variables). This means that the MA form for the SVAR is $z_t = C(L)\eta_t$ and it can be partitioned as

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} C_{11}(L) & 0 \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}.$$

Applying a Q matrix to the shocks η_t produces $z_t = C(L)Q'Q\eta_t = C^*(L)\eta_t^*$, which in partitioned form will be (writing Q' as F)

$$\begin{aligned} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} &= \begin{bmatrix} C_{11}(L) & 0 \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \eta_{1t}^* \\ \eta_{2t}^* \end{bmatrix} \\ &= \begin{bmatrix} C_{11}(L)F_{11} & C_{11}(L)F_{12} \\ C_{21}(L)F_{11} + C_{22}(L)F_{21} & C_{21}(L)F_{12} + C_{22}(L)F_{22} \end{bmatrix} \begin{bmatrix} \eta_{1t}^* \\ \eta_{2t}^* \end{bmatrix}. \end{aligned}$$

To ensure that η_{1t}^* corresponds to foreign shocks, and that foreign variables are not affected by domestic shocks at any lags, it is necessary to have $F_{12} = 0$. Consequently, since $F = Q'$ this means that $Q_{21} = 0$. But $Q'Q = I_n$ so we must have

$$\begin{bmatrix} F_{11} & 0 \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ 0 & Q_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

giving $F_{21}Q_{11} = 0$, i.e. $F_{21} = 0$. But $F_{21} = Q_{12}$ so this means Q must have the form $\begin{bmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{bmatrix}$, i.e. it is necessary to separately combine the foreign and domestic base shocks. Although we don't have $Q_{11}'Q_{22} = 0$ the new shocks $\eta_{1t}^* = Q_{11}\eta_{1t}$ and $\eta_{2t}^* = Q_{22}\eta_{2t}$ remain uncorrelated, because $E(Q_{11}\eta_{1t}\eta_{2t}'Q_{22}') = 0$, owing to $E(\eta_{1t}\eta_{2t}') = 0$.

5.6 Standard Errors for Sign Restricted Impulses

5.6.1 The SRR Method

Let $\hat{C}_j$ be the impulse responses at the $j'th$ lag for one standard deviation shocks from a recursive model. These imply that

$$z_t = \hat{C}_0\eta_t + \hat{C}_1\eta_{t-1} + ...,$$

where η_t are the standardized recursive shocks, i.e. the base shocks. Then

$$\begin{aligned} z_t &= \hat{C}_0 Q'Q\eta_t + \hat{C}_1 Q'Q\eta_{t-1} + ... \\ &= \hat{C}_0 Q'\eta_t^* + \hat{C}_1 Q'\eta_{t-1}^* + ..., \end{aligned}$$

and η_t^* are the sign-restricted shocks. We therefore have that $\hat{C}_j^* = \hat{C}_j Q'$ so that $vec(\hat{C}_j^*) = (Q \otimes I)vec(\hat{C}_j)$.

Now assume that $vec(\hat{C}_j)$ is normal with mean $vec(\bar{C}_j)$ and variance V (at least in large samples). The mean need not be equal to the true impulse responses since the recursive model that begins the process is most likely misspecified. Hence the mean of $vec(\hat{C}_j^*)$ will be $(Q \otimes I)E[vec(\hat{C}_j)]$ while the variance of $\hat{C}_j^*$ will be

$$\begin{aligned} var(vec(\hat{C}_j^*)) &= var\{vec(\hat{C}_j^*) - (Q \otimes I)vec(\bar{C}_j))\} \\ &= var\{(Q \otimes I)(vec(\hat{C}_j) - vec(\bar{C}_j))\} \\ &= (Q \otimes I)var(\hat{C}_j)(Q' \otimes I) \end{aligned}$$

Consequently the standard errors of the impulses vary according to the model as summarized by Q. There is also a common component which is the variance of the standardized recursive model shocks that initiated the process. The bias will also be different.

5.6.2 The SRC method

In the case of the SRC method the standard errors will reflect the method used to capture the estimable parameters. It is possible to use any method that will estimate the parameters of a structural system, e.g. FIML, IV, Bayesian methods. The standard errors found will vary from realization to realization. Once a model is selected, then standard errors follow immediately.

5.7 Summary

Sign restrictions look attractive as they are acceptable to many investigators, but *weak information* gives *weak results*. Often the results from sign restricted SVARs have been presented as if the results are strong. In this chapter we have argued that this is illusory. There are many unresolved problems with the methodology. Getting a single set of impulse response functions is a key one. To do this one needs to impose some *extra* information and that will be context and institution dependent. In general one needs to think carefully about the modeling process, and it seems doubtful that the methodology can be automated. Combinations of parametric and sign restrictions would seem to be the best use of this type of restriction rather than to just use it to the exclusion of parametric methods.

Chapter 6

Modeling SVARs with Permanent and Transitory Shocks

6.1 Introduction

To date it has been assumed that all the variables that we are working with are covariance stationary. But there is an increasing recognition that many economic variables cannot be described in this way, and the nature of variables can affect both the way an SVAR analysis is performed and the type of restrictions that can be applied. Section 2 of this chapter looks at the nature of variables, making a distinction between variables according to whether they are integrated of order one (i.e., I(1)) or zero (i.e., I(0)). A variable that is I(1) is said to have a stochastic trend and so can be thought of as non-stationary, while an I(0) variable might be taken to be stationary.

Sometimes we will use this language of stationarity, although it is not a rigorous distinction. An I(1) variable typically has a permanent component but an I(0) variable only has a transitory component.[1] Section 2 of the chapter looks at the nature of these variables and the two types of shocks that can be present in SVAR systems that have some I(1) variables in them. When faced with the distinction between permanent and transitory components one solution has been to extract a permanent component from the I(1) variables and use what is left, the transitory component, in place of the variables themselves. Perhaps the most favored approach is to use Hodrick-Prescott filtered data in the SVAR. Consequently, Section 3 explains why this is not a valid methodology. For the remainder of this chapter the original (unfiltered) data will be used in an SVAR.

Section 4 then goes through a series of examples to show how the methods of

[1]An exception is an I(1) variable driven by MA(1) innovations with a negative unit root (i.e., $\epsilon_t - \epsilon_{t-1}$).

previous chapters need to be modified to deal with I(1) and I(0) variables. These examples all deal with the situation when there is no cointegration between the non-stationary variables. Chapter 7 will describe the adjustments needed if there is cointegration. The key feature of this chapter is that variables can be either stationary or non-stationary, but the SVAR needs to be set up in such a way as to respect the distinction. The applications in Section 4 start with a simple two variable money/income model where both variables are non-stationary, moves on to a two-variable model with a mixture of these features - the Blanchard and Quah (1989) model - and then to a four variable model of oil price shocks developed by Peersman (2005). Lastly, we return to the small macro model. Previously its data had been adjusted so as to ensure that all variables were stationary, but now we work with the unadjusted GDP data and treat it as being I(1). It becomes apparent that this change makes a big difference to the impulse responses for monetary shocks.

6.2 Variables and Shocks

Economic series were originally viewed as stationary around a *deterministic trend,* i.e.,

$$z_t = y_t - \phi - \psi t = b_1 z_{t-1} + e_t \ (b_1 < 1). \tag{6.1}$$

Estimates of ϕ and ψ could be found by regressing y_t on a constant and time and residuals $\hat{z}_t$ could be constructed. This approach accounted for the dual facts that (i) visually it seemed clear that there were consistent upward or downward movements in many series and (ii) the deviations from these upward or downward movements were persistent. If y_t was the log of GDP then $\hat{z}_t$ became known as an output gap.

The new view of time series starting in the late 1950s was that $b_1 = 1$ (see Quenouille, 1957). In order to contrast the two cases we will put $\phi = 0, \psi = 0$ so that $z_t = y_t$. Then the AR(1) series $y_t = b_1 y_{t-1} + e_t$ has a *unit root* if $b_1 = 1$, and the series y_t is said to be (integrated of order 1). This process is non-stationary. There are series that are non-stationary but not integrated, e.g. some of the fractionally integrated class of processes, but these will be ignored.[2]

Once $I(1)$ processes emerged it was recognized that some extra concepts were needed. One of these was to observe that a unit root process meant that the variance of z_t was infinite. A good way of seeing this is to derive the variance of y_t given that $y_0 = 0$. This is $t\sigma_e^2$, so that the variance rises consistently with time and eventually becomes infinite. Because the variance depends on t the I(1) series is said to have a *stochastic* trend, as compared to the *deterministic* trend of (6.1), where the word "trend" here is being used in two different ways. A way of differentiating them is to ask what is the probability that y_t would return to y_0 as $t \to \infty$. For (6.1) this probability goes to zero. However, when $\phi = 0, \psi = 0, b_1 = 1$, there is a pure random walk process that always returns to where it started, although the time between returns lengthens, owing to the

[2] See Baillie (1996) for a review of long memory processes.

rise in the variance. Thus much of what is seen in graphs of series exhibiting upward and downward movements is about whether $\phi \neq 0$, not whether there is a unit root process.

A second addition to the concepts needed for analysis related to the nature of the shocks. Consider the experiment of raising e_t by one unit at t and resetting it to its initial value from $t+1$ onward. Its impact (in a pure random walk) is for y_{t+j} to rise by one for all $j > 0$, and so the effect of this change is *permanent.* If however $|b_1| < 1$, then the series is $I(0)$ and, performing the same experiment, y_{t+j} rises by b_1^j. This tends to zero as $j \to \infty$ (provided $|b_1| < 1$) and, because it dies out, the *effect* is *transitory.* In the first case e_t is called a *permanent shock* and, in the second, it is a *transitory shock.* Notice that this is about the *effect* of the shock, not its *nature.* In both cases the shock is a non-integrated (stationary) process. Related to these distinctions pertaining to the effect of the shocks was that a series y_t could be decomposed into a permanent and a transitory component. If a series was $I(0)$ then it only had a transitory component.

We will need to extend these definitions regarding shocks to handle more than one series. Let us assume that there are three series. Therefore there will be three shocks and a 3×3 matrix C showing the *long-run effects* of the three shocks upon the three variables. The long-run effects of the $k'th$ shock on the $l'th$ variable will be $\sum_{\lim j \to \infty} \frac{\partial y_{lt+j}}{\partial e_{kt}}$, so the C matrix will have variables arranged in the rows and shocks in the columns. Suppose we begin by assuming that all three variables are $I(1)$ and that the matrix looks like

$$\mathrm{C} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & 0 & 0 \end{bmatrix} \end{array},$$

where * means that the effect is non-zero. Thus in this case the first shock(s_1) has a non-zero long run effect on all the three variables in the system. The second shock (s_2) affects the first two variables in the system and the last shock(s_3) does not affect any of the variables in the long run. Thus the first two shocks are permanent, while the last shock is transitory since it has a zero long-run effect on *all* the variables. Notice that a permanent shock can have zero long-run effects on *some* $I(1)$ variables *but not on all of them.* The rank of the matrix C gives the number of permanent shocks, and in this case it is clearly two, implying that there is only one transitory shock.

Now let us suppose that instead of there being three $I(1)$ variables we have two $I(1)$ variables and one $I(0)$ variable. Let the $I(0)$ variable be the third one (row 3 of the assumed C matrix). In this instance suppose that the matrix C looks like

$$\mathrm{C} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Because the third variable is $I(0)$ the last row has only zero elements. Apart from that the matrix shows that the first shock has a permanent effect on the first variable, the second shock is transitory since it has zero long-run effects on the I(1) variables, and the third shock has a permanent effect on the second variable. Because the rank of this matrix is two there will be two permanent shocks. The importance of this case is to emphasize that the nature of variables and the nature of shocks can be quite different. We will encounter these cases in the applications that follow.

6.3 Why Can't We Use Transitory Components of $I(1)$ Variables in SVARs?

The issue investigated in this section involves filtering an $I(1)$ series in some way so as to extract the transitory component and then using this filtered series in a VAR. Because there is no unique way of performing a permanent/transitory decomposition each method adds on some extra constraint so as to get a single division into the components. It is useful to look at one of these - the Beveridge-Nelson (BN) decomposition - as the lessons learnt from it are instructive.

Suppose we have n variables that are $I(1)$ and not co-integrated. Then the Beveridge and Nelson decomposition defined the permanent component of y_t as

$$y_t^{BN,P} = \lim_{T\to\infty} E_t(y_T) = y_t + E_t \sum_{j=1}^{\infty} \Delta y_{t+j},$$

from which it is necessary to describe a process for Δy_t so as to compute $y_t^{BN,P}$. Suppose it is a VAR(1), $\Delta y_t = \Gamma_1 \Delta y_{t-1} + \varepsilon_t$. It follows that the transitory component $(y_t - y_t^{BN,P})$ would be

$$y_t^{BN,T} = y_t - y_t^{BN,P} = -E_t \sum_{j=1}^{\infty} \Delta y_{t+j} = -\Gamma_1 (I_n - \Gamma_1)^{-1} \Delta y_t.$$

Extending this to Δy_t being a VAR(p) with coefficients $\Gamma_1, ..., \Gamma_p$, one would get $y_t^{BN,T} = -\sum_{j=0}^{p-1} \Phi_j \Delta y_{t-j}$ where Φ_j will be functions of $\mathit{\Gamma}_1$,....,$\mathit{\Gamma}_p$.

This analysis points to the fact that, if Δy_t follows a VAR(p) then the filtered (transitory) component of y_t would weight together $\Delta y_t, .., \Delta y_{t-p+1}$. Consequently suppose an SVAR(2) is assumed of the form

$$A_0 \Delta y_t = A_1 \Delta y_{t-1} + A_2 \Delta y_{t-2} + \varepsilon_t,$$

that is $A(L)\Delta y_t = \varepsilon_t$, where $A(L) = (A_0 - A_1 L - A_2 L^2)$. So $\Delta y_t = A(L)^{-1} \varepsilon_t$ and, because $y_t^{BN,T} = \Phi_0 \Delta y_t + \Phi_1 \Delta y_{t-1}$, we have

$$y_t^{BN,T} = \Phi_0 A(L)^{-1} \varepsilon_t + \Phi_1 A(L)^{-1} \varepsilon_{t-1},$$

showing that the process for the transitory component *is not* a VAR, except in the special case where y_t is a scalar and Δy_t is an AR(1). In this case Φ_0

is a scalar, $\Phi_1 = 0$, and so we can write $A(L)y_t^{BN,T} = \Phi_0\varepsilon_t$. For y_t being of higher dimension $\Phi_0 A(L)^{-1}$ does not commute to $A(L)^{-1}\Phi_0$. Hence, using the transitory component from a BN filter in a finite order VAR would be in error.

Now there are other ways of extracting a transitory component which involve averaging the y_t to eliminate a permanent component. These are filters such as Hodrick-Prescott and the Band-Pass class. They all have the following filter structure

$$\begin{aligned} y_t^P &= \sum_{j=0}^{m} \omega_{\pm j} y_{t\pm j} \\ &= \sum_{j=0}^{m} \omega_{\pm j} y_t + \sum_{j=1}^{m} \omega_j \Delta_j y_{t+j} - \sum_{j=1}^{m} \omega_{-j} \Delta_j y_t, \end{aligned}$$

where $\Delta_k y_t = y_t - y_{t-k}$.[3] Most filters are symmetric so that $\omega_{-j} = \omega_j$ and so the transitory component would be

$$y_t^T = y_t - y_t^p = (1 - \sum_{j=0}^{m} \omega_{\pm j}) y_t - \sum_{j=1}^{m} \omega_j \Delta_j y_t + \sum_{j=1}^{m} \omega_j \Delta_j y_{t+j}.$$

It will be necessary for $1 - \sum_{j=0}^{m} w_{\pm j} = 0$, otherwise y_t^T would be non-stationary (as y_t is). This means that the transitory component will be

$$y_t^T = \sum_{j=1}^{m} \omega_j \Delta_j y_{t+j} - \sum_{j=1}^{m} \omega_j \Delta_j y_t.$$

The ω_j then come from applying many criteria pertaining to the nature of the permanent and transitory components. A Band-Pass filter focuses upon frequencies of the spectrum. The Hodrick-Prescott filter chooses them to make the permanent component smooth.[4] So each of these provides a different weighted average of the current, past and future growth rates.

Again, the same issue arises as with the BN filter - the process in the filtered series would not be a VAR. But the situation is worse here since, unless $m = 0$, there will always be an MA structure to the transitory component and, whenever $m > 0$, the filtered data will depend on *future shocks,* because the HP and Band-Pass class are two-sided filters, compared to the one-sided nature of BN. Clearly it is very unsatisfactory to use two sided filters like this in any regression. Doing so will produce inconsistent estimators of coefficients.

[3] Note that $\Delta_j y_t = y_t - y_{t-j} = \Delta y_t + \Delta y_{t-1} + .. + \Delta y_{t-j+1}$

[4] When $\lambda = 1600$ and the HP filter is applied to quarterly data, $m = 12$ gives a reasonable approximation to the HP filtered data on the transitory component.

6.4 SVARs with Non-Cointegrated $I(1)$ and $I(0)$ Variables

6.4.1 A Two Variable System in $I(1)$ Variables

We assume that there are two non-cointegrated $I(1)$ variables ζ_{1t} and ζ_{2t} (these will be taken to be the logs of variables expressed in levels). Because there is no cointegration it will be necessary to work with first differences, i.e. the SVAR will be expressed in terms of $z_{1t} = \Delta\zeta_{1t},\ z_{2t} = \Delta\zeta_{2t}$. For illustrative purposes it is taken to be an SVAR(1):

$$A_0 z_t = A_1 z_{t-1} + \varepsilon_t.$$

To flesh this out let $\zeta_{1t} =$ log output and $\zeta_{2t} =$ log money supply. Then $z_{1t} =$ output growth and $z_{2t} =$ money supply growth. Because there are two $I(1)$ variables and no cointegration there must be two permanent shocks. The structural impulse responses will come from $z_t = C(L)\varepsilon_t$.

Now let us look at the implications for the impulse responses of the $I(1)$ nature of the series. By definition

$$\begin{aligned}
\zeta_{t+M} &= \zeta_{t-1} + \sum_{k=0}^{M} \Delta\zeta_{t+k} = \zeta_{t-1} + \sum_{k=0}^{M} z_{t+k} \\
\text{so} \quad & \frac{\partial \zeta_{t+M}}{\partial \varepsilon_t} = \sum_{k=0}^{M} \frac{\partial z_{t+k}}{\partial \varepsilon_t} = \sum_{k=0}^{M} C_k \\
\Longrightarrow \quad & \lim_{M\to\infty} \frac{\partial \zeta_{t+M}}{\partial \varepsilon_t} = \sum_{k=0}^{\infty} C_k
\end{aligned}$$

Because $C(L) = C_0 + C_1 L + C_2 L^2 + ...$, a shorthand for $\sum_{k=0}^{\infty} C_k$ is $C(1)$, and, in line with the terminology used in the introduction, this will be termed the matrix of *"long-run responses"*. It shows the effects of a shock at t on the *levels of* ζ_t at infinity. Consequently, it can therefore be used to define the *long-run* effects of shocks. It also serves to define a transitory shock ε_{kt} as one for which the $k'th$ column of $C(1)$ is all zeros. If there exists any non-zero element in the $k'th$ column of $C(1)$, it means that the $k'th$ shock is *permanent*. It needs to be stressed once again that a permanent shock *need not affect all $I(1)$ variables, just one.*

Now consider the case where the second shock has a zero long-run effect on ζ_{1t}. This can be summarized by the *long-run response* matrix

$$C(1) = \begin{bmatrix} c_{11}(1) & 0 \\ c_{12}(1) & c_{22}(1) \end{bmatrix}, \tag{6.2}$$

which sets the $c_{12}(1)$ element to 0. Note that with this assumption the second shock is *permanent* because $c_{22}(1) \neq 0$.

Table 6.1: Estimating the Money-Output Growth Model with Long-Run Restrictions using EViews

File → Open →EViews Workfile
Locate ***gdp_m2.wf1*** *and open it*
Object→New Object→Matrix-Vector-Coef
Choose *matrix, 2 rows, 2 columns* and fill in the 2x2 matrix
as $\begin{bmatrix} NA & 0 \\ NA & NA \end{bmatrix}$, naming it $C1$ in the workfile
Quick →Estimate VAR
Endogenous Variables s_*dgdp s_dm2*
Lag Intervals for Endogenous 1 1
Exogenous Variables *c*
Estimation Sample 1981q3 2000q1
Proc →Estimate Structural Factorization→Matrix
Choose *matrix* opting for *long-run pattern* and putting
C1 in as the name and then ***OK***
Impulse→Impulse Definition →Structural Decomposition
These are the responses of d*gdp to dm2.* If you want *levels of log gdp, log m2*
choose instead ***Accumulated Responses*** at the ***Impulse Definition*** box

The form of $C(1)$ is crucial to the chapter, and the key to handling $I(1)$ processes is in determining what the zeros in $C(1)$ imply regarding parametric restrictions on the SVAR representing z_{jt}.

6.4.1.1 An EViews Application of the Two I(1) Variable Model

The two variables ζ_{jt} will be the log of GDP and the log of real M2 balances. The restriction is that just described, namely the second permanent shock (i.e., money) has a zero long-run effect on output (i.e., log GDP). The SVAR is expressed in terms of $z_{1t} = \Delta\zeta_{1t}$ and $z_{2t} = \Delta\zeta_{2t}$, i.e. the growth rates in GDP and money (these are called *s_dgdp, s_dm2* in the data set *gdp_m2.wf1*). The matrix $C = C(1)$ needs to be specified and it has the form in (6.2). Table 6.1 gives the EViews commands to estimate this model and the resulting output and accumulated impulse response functions are shown in Figures 6.1 and 6.2 respectively.[5] Note that EViews does not provide standard errors for the impulse responses using this approach. That can be done by using an alternative approach to estimation originally proposed in Shapiro and Watson (1988).

[5] When EViews estimates an SVAR with long-run restrictions it works with $e_t = A_0^{-1}\Gamma u_t$, where Γ is a diagonal matrix containing the standard deviation of the structural errors. The procedure for extracting the final estimate of A_0 is explained in footnote 6 in Chapter 4. Using this approach yields $F^{-1}A_0 = \begin{bmatrix} 0.546303 & -0.254799 \\ 0.301918 & 0.873391 \end{bmatrix}$, impying that $A_0 = \begin{bmatrix} 1 & 0.291735 \\ -0.55266 & 1 \end{bmatrix}$ after renormalization.

Var: GDPM Workfile: UNTITLED::gdp_m2\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 02/02/16 Time: 17:43
Sample: 1981Q3 2000Q1
Included observations: 75
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 5 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: long-run pattern matrix
Long-run response pattern:

C(1)	0
C(2)	C(3)

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.052866	0.085966	12.24745	0.0000
C(2)	0.508288	0.203255	2.500744	0.0124
C(3)	1.723155	0.140695	12.24745	0.0000

Log likelihood -168.5552

Estimated A matrix:

1.000000	0.000000
0.000000	1.000000

Estimated B matrix:

0.546303	-0.254799
0.301918	0.873391

Figure 6.1: SVAR Results for the Money/GDP Model: Zero Long-Run Effect of Money on GDP

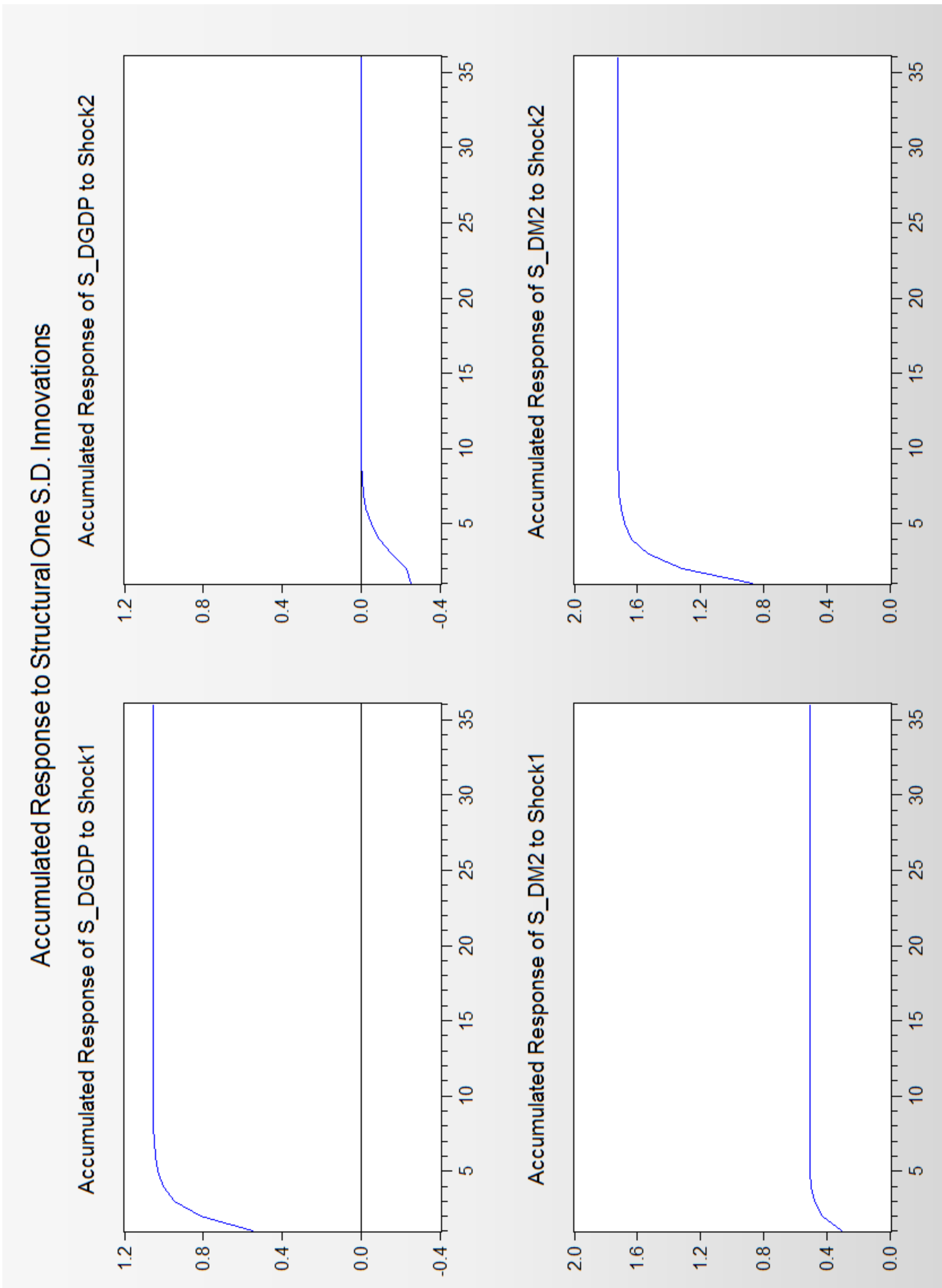


Figure 6.2: Impulse Responses from the Money-Output Model with Zero Long-Run Effect of Money on GDP

6.4.1.2 An Alternative EViews Application of the Two I(1) Variable Model

Shapiro and Watson (1988) highlight that the restrictions on $C(1)$ imply a specific form for the SVAR that can be estimated directly, provided the restrictions are imposed. The SVAR is $A(L)z_t = \varepsilon_t$, the underlying VAR is $B(L)z_t = e_t$, and the MA form for the structural errors is $z_t = C(L)\varepsilon_t$. Hence

$$\begin{aligned} C(L) &= A(L)^{-1} \Longrightarrow C(L)A(L) = I_n \\ &\Longrightarrow C(1)A(1) = I_n \end{aligned}$$

Because the two variable SVAR(1) system is

$$z_{1t} = a_{12}^0 z_{2t} + a_{11}^1 z_{1t-1} + a_{12}^1 z_{2t-1} + \varepsilon_{1t} \tag{6.3}$$

$$z_{2t} = a_{21}^0 z_{1t} + a_{21}^1 z_{1t-1} + a_{22}^1 z_{2t-1} + \varepsilon_{2t}, \tag{6.4}$$

we have $A(L) = A_0 - A_1 L = \begin{bmatrix} 1 - a_{11}^1 L & -a_{12}^0 - a_{12}^1 L \\ -a_{21}^0 - a_{21}^1 L & 1 - a_{22}^1 L \end{bmatrix}$. Consequently, with $C(1)$ defined as in (6.2), $C(1)A(1) = I_n$ can be written as

$$\begin{bmatrix} c_{11}(1) & 0 \\ c_{12}(1) & c_{22}(1) \end{bmatrix} \begin{bmatrix} 1 - a_{11}^1 & -a_{12}^0 - a_{12}^1 \\ -a_{21}^0 - a_{21}^1 & 1 - a_{22}^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

which means that $c_{11}(1)(-a_{12}^0 - a_{12}^1) = 0$. Because $c_{11}(1) \neq 0$ ($C(1)$ would be singular if it wasn't) it follows that $a_{12}^0 + a_{12}^1 = 0$, i.e. $a_{12}^0 = -a_{12}^1$. Imposing this restriction upon (6.3) we get

$$z_{1t} = a_{11}^1 z_{1t-1} + a_{12}^0 \Delta z_{2t} + \varepsilon_{1t}. \tag{6.5}$$

Hence long-run restrictions result in parametric restrictions *between* the elements in A_0 and A_1, implying that z_{2t-1} does not appear in (6.5), and it can therefore be used as an instrument for Δz_{2t}. Once (6.5) is estimated $\hat{\varepsilon}_{1t}, z_{1t-1}$ and z_{2t-1} can be used as instruments to estimate (6.4).[6]

Once the point estimates for a_{12}^0 and a_{21}^0 (of -.291735 and .552657 respectively) are found, the restricted SVAR can be defined through the A matrix used in EViews, and the remaining coefficients in the SVAR can be estimated.[7] The accumulated impulse responses are shown in Figure 6.3. They are identical for both approaches to estimation; however, an advantage of proceeding with the Shapiro and Watson method is that standard errors for the impulse responses will be supplied by EViews.

Rather than provide menu instructions as in Table 6.2, in what follows we will provide EViews command line code that produces the same output - see

[6] One can apply this to a $VAR(p)$, in which case the first equation would have as regressors $\Delta z_{2t}, \Delta z_{2t-1}, .., \Delta z_{t-p+1}$, and z_{2t-p} would be used as the instrument for Δz_{2t}. The result will be that the sum of the parameter estimates corresponding to z_{2t} will be zero (i.e., $\sum_{j=0}^{p}[A_j]_{12} = 0$).

[7] The values of a_{12}^0 and a_{21}^0 found with the standard EViews SVAR approach (namely Table 6.1) and the IV approach as in Table 6.2 are identical.

Table 6.2: An IV Method for Fitting the Money-Output Growth Model with Long-Run Restrictions

File → Open →EViews Workfile
Go to directory where gdp_m2.wf1 is and click on it
Quick →Estimate Equation
choose 2SLS for the instrument option - first equation
Equation Specification *s_dgdp s_dgdp(-1) d(s_dm2) c*
Instrument List *s_dgdp(-1) s_dm2(-1)*
Make sure constant is in instrument list (check box)
Estimation Sample 1981q3 2000q1
Proc→ Make Residual Series → Name *eps1*
Estimate
Equation Specification *s_dm2 s_dgdp s_dgdp(-1) s_dm2(-1) c*
Instrument List *s_dgdp(-1) s_dm2(-1) eps1*
Make sure constant is in instrument list (check box)
Quick →Estimate VAR
Endogenous Variables s_*dgdp s_dm2*
Lag Intervals for Endogenous *1 1*
Exogenous Variables c
Proc →Estimate Structural Factorization
@e1=-0.291735*@e2+c(1)*@u1
@e2=c(3)*@e1+c(2)*@u2
Impulse→ Impulse Definition → Structural Decomposition

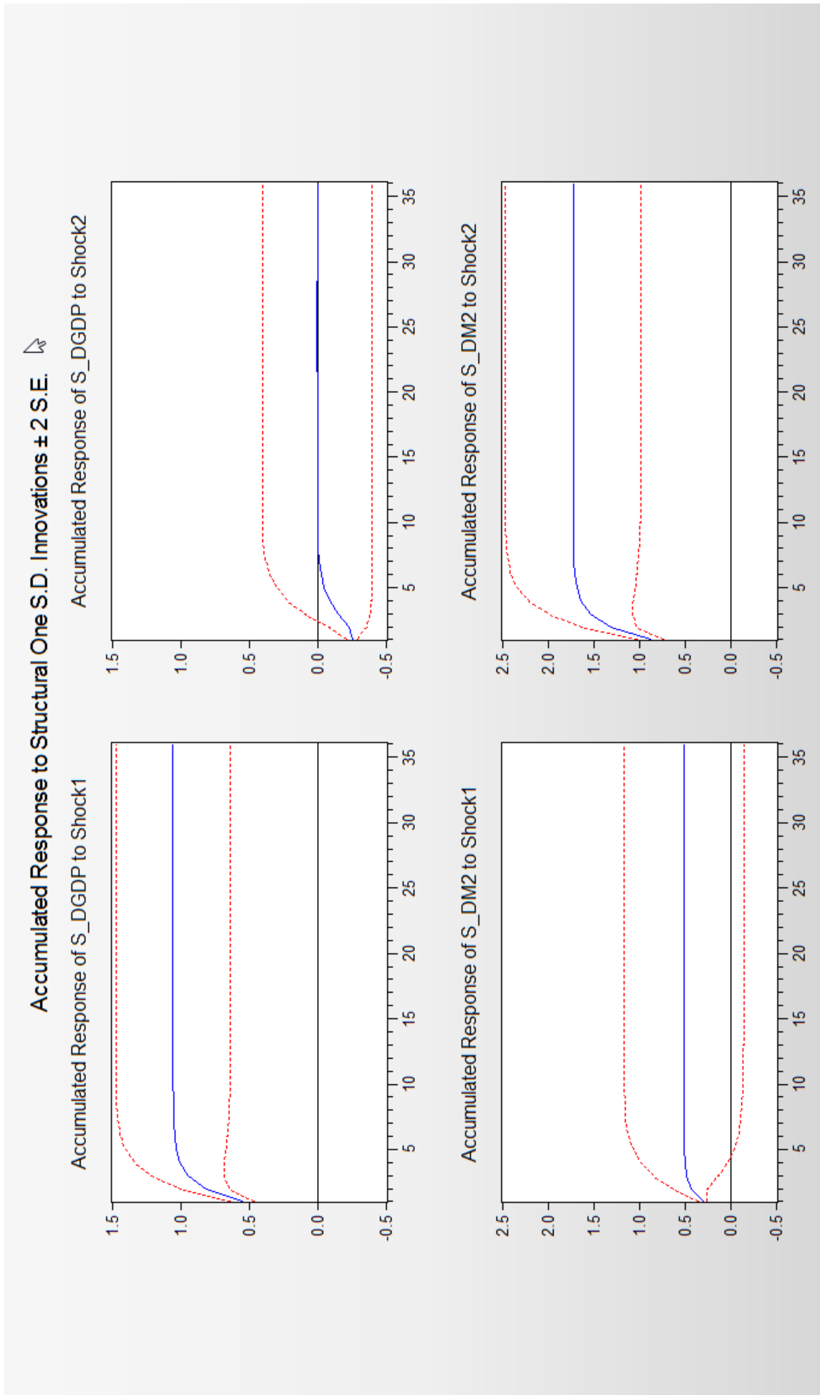
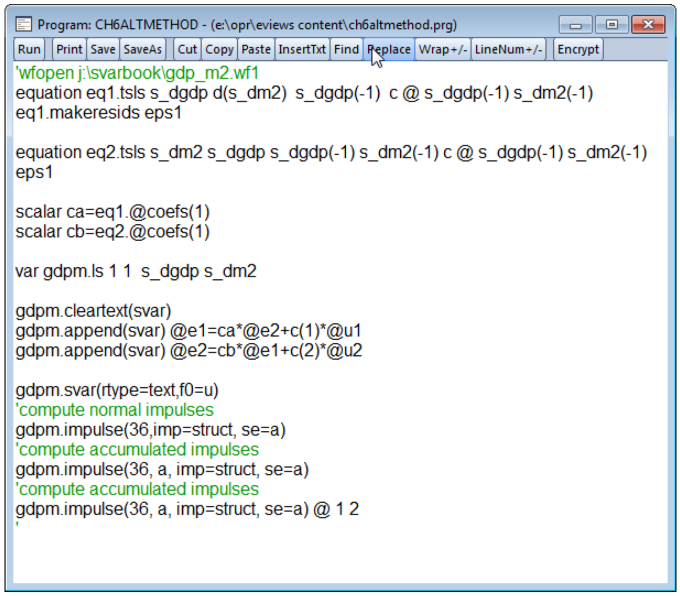


Figure 6.3: Accumulated Impulse Responses from the Money-Output Model with Zero Long-Run Effect of Money on GDP




```
'wfopen j:\svarbook\gdp_m2.wf1
equation eq1.tsls s_dgdp d(s_dm2)  s_dgdp(-1)  c @ s_dgdp(-1) s_dm2(-1)
eq1.makeresids eps1

equation eq2.tsls s_dm2 s_dgdp s_dgdp(-1) s_dm2(-1) c @ s_dgdp(-1) s_dm2(-1)
eps1

scalar ca=eq1.@coefs(1)
scalar cb=eq2.@coefs(1)

var gdpm.ls 1 1  s_dgdp s_dm2

gdpm.cleartext(svar)
gdpm.append(svar) @e1=ca*@e2+c(1)*@u1
gdpm.append(svar) @e2=cb*@e1+c(2)*@u2

gdpm.svar(rtype=text,f0=u)
'compute normal impulses
gdpm.impulse(36,imp=struct, se=a)
'compute accumulated impulses
gdpm.impulse(36, a, imp=struct, se=a)
'compute accumulated impulses
gdpm.impulse(36, a, imp=struct, se=a) @ 1 2
'
```

Figure 6.4: EViews Program *ch6altmethod.prg* to Impose a Long-Run Restriction in the Money-Output Model using IV Methods

Figure 6.4 for an EViews program *(ch6altmethod.prg)* that reproduces the work described in Table 6.1.

A different way to proceed in EViews is to create the two equation system using the ***Proc→Make System→By Variable*** option that becomes available after running a VAR using *s_dgdp and s_dm2.* The resulting system object can then be edited so as use instrumental variables. Calling this system GDPMIV, it can be found in the workfile *gdp_m2.wf1* and contains the code:

*S_DGDP = C(1)*S_DGDP(-1) + C(2)*D(S_DM2)+ C(3) @ S_DGDP(-1) S_DM2(-1) C*

*S_DM2 = C(4)*S_DGDP(-1) + C(5)*S_DM2(-1) +C(6)*S_DGDP + C(7) @ S_DGDP(-1) S_DM2(-1) C (S_DGDP - C(1)*S_DGDP(-1) - C(2)*D(S_DM2)- C(3))*

Notice that the last instrument (i.e., the estimated residuals of the first equation) appears after the @ sign in the second equation, and references the estimated coefficients (i.e., C(1), C(2) and C(3)) elements explicitly to ensure that the residuals used in the second equation as instruments are equal to the implied residuals of the first equation. Then choosing ***Estimate→Two Stage Least Squares*** one gets the same parameter estimates as from the previous program. To also get the impulse responses with the system approach we run the program *gdpmsystem.prg* shown in Figure 6.5.

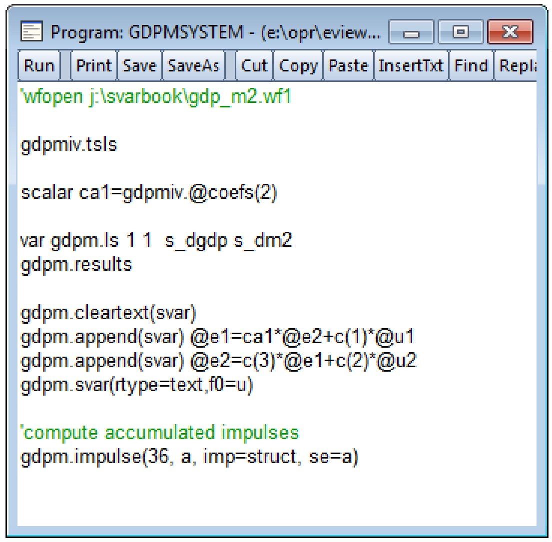


Figure 6.5: EViews Program gdpmsystem.prg to Produce Impulse Response Functions for the Money-Output Model

As explained in Chapter 4, an equivalent approach to estimating the model is to use a FIML estimator together with the restrictions needed to ensure that the long-run response of z_{1t} to a shock in z_{2t} is zero. The first restriction is that the residual covariance matrix is diagonal, thereby ensuring that the model is structural. The second is that the sum of the coefficients associated with z_{2t} (i.e., contemporaneous and lagged) in the first equation of the structural VAR sum to zero (see Equation 6.5). The required EViews code to impose the adding up constraint is shown in Figure 6.6. Notice that for the first equation the sum of the coefficients on S_DM2 and its lag (i.e., SM2_DM2(-1)) is zero. The system is exactly identified because of this constraint.

Estimating the model with FIML and the diagonal covariance option yields the results presented in Figure 6.7. The parameter estimates for C(2) and C(6) match those obtained using the SVAR routine in EViews.[8]

6.4.2 A Two Variable System with a Permanent and Transitory Shock - the Blanchard and Quah Application in EViews

Blanchard and Quah (BQ) (1989) dealt with a case where there were two series, one of which was $I(1)$ and the other $I(0)$, and the shock in the structural

[8]The program *gdp_m2_mle.prg* in the MLE sub-directory uses the *optimize*() routine in EViews to implement the FIML estimator for this model by maximizing a user-defined likelihood function.

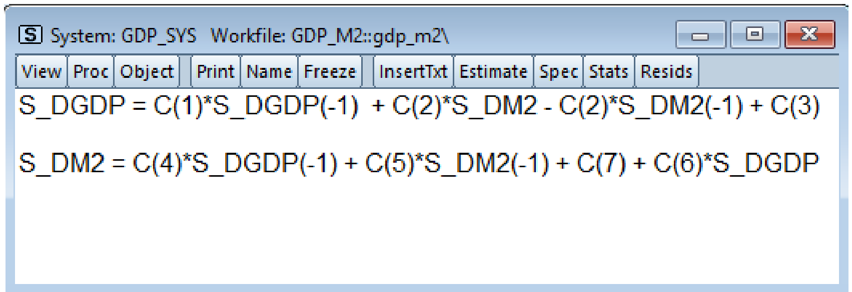


Figure 6.6: EViews SYSTEM Object to Estimate Money/GDP Model with A Zero Long-Run Restriction

equation for the $I(0)$ variable was transitory. To be more precise, in their case ζ_{1t} is $I(1)$ and ζ_{2t} is $I(0)$, where ζ_{1t} = log GNP and ζ_{2t} = the detrended unemployment rate (u_t). They estimated a SVAR in $z_{1t} = \Delta\zeta_{1t}$ and $z_{2t} = \zeta_{2t}$. Because of their assumptions, there must be one permanent shock and one transitory shock. The economic rationale given for this is that a demand shock typically has a transitory effect on output, i.e. it has a zero long-run effect on GNP, while a supply shock has a permanent effect on GNP.

Again we have the moving average representation $z_t = C(L)\varepsilon_t$. Recall that the values of $C(1)$ define the long-run responses of ζ_{1t} (first row of C(1)) and ζ_{2t} (second row of C(1)) to the shocks in the system. The restriction relating to the demand shock implies that $c_{12}(1) = 0$. Now the first equation of the SVAR(1) is

$$\Delta\zeta_{1t} = a_{12}^0\zeta_{2t} + a_{11}^1\Delta\zeta_{1t-1} + a_{12}^1\zeta_{2t-1} + \varepsilon_{1t}.$$

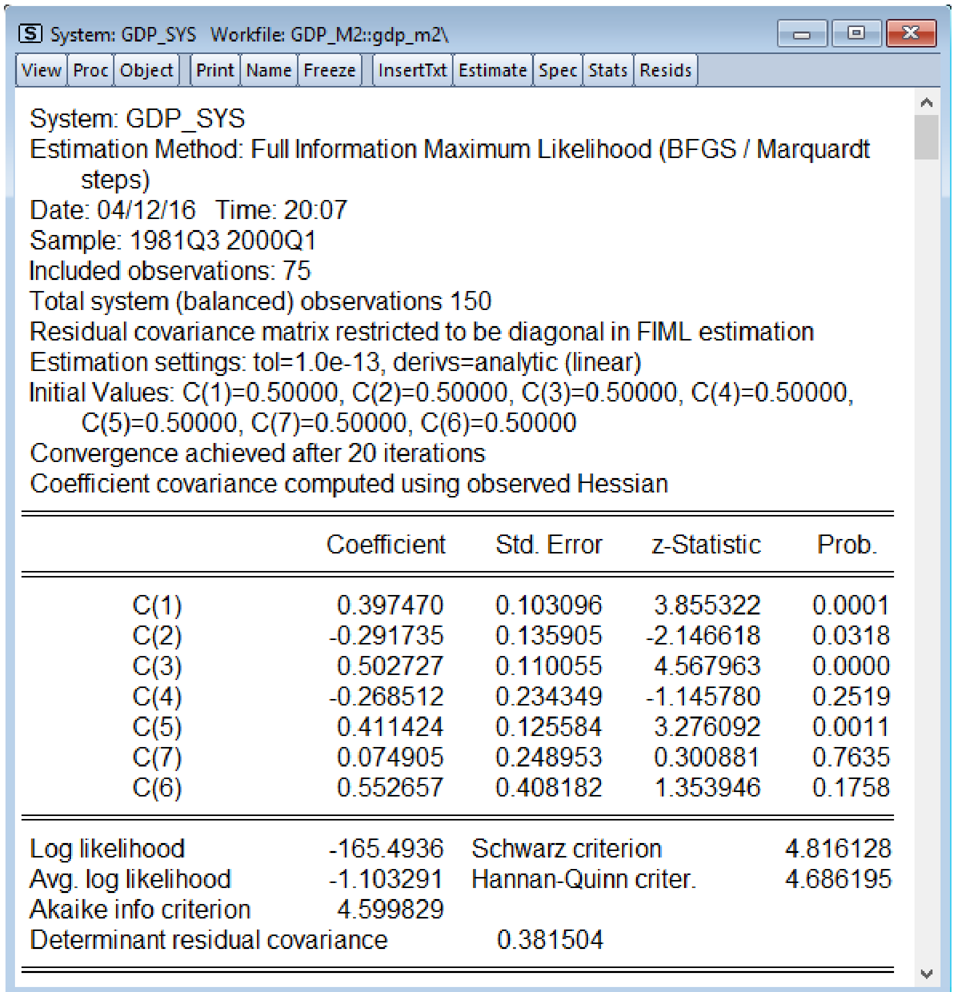
Just as in the previous application, after imposing the restriction $c_{12}(1) = 0$, this equation becomes

$$\Delta\zeta_{1t} = a_{12}^0\Delta\zeta_{2t} + a_{11}^1\Delta\zeta_{1t-1} + \varepsilon_{1t},$$

and ζ_{2t-1} can be used as an instrument for $\Delta\zeta_{2t}$. So this is just like the Money/GDP case considered in the previous section, and the same estimation procedures can be used. In *bqdata.wf1* the variables $\Delta\zeta_{1t}$ and ζ_{2t} are named dya_t and u_t respectively with Δu_t being du_t.

The major difference between this and the applications of the preceding sub-section is that Blanchard and Quah use a SVAR(8) rather than a SVAR(1). However, the method of handling a VAR(p) was discussed earlier in footnote 6. Program *bq.prg* in Figure 6.8 contains the required command line code, Figure 6.9 shows the parameter estimates, and Figure 6.10 the impulse responses.

Similarly, the system object code to replicate Blanhard-Quah's application using EViews' FIML estimator is given in Figure 6.11. Note that the coefficient on the unemployment rate, u, in the first (output) equation is constrained to equal the negative of the sum of the lagged coefficients on the unemployment



System: GDP_SYS Workfile: GDP_M2::gdp_m2\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: GDP_SYS
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Date: 04/12/16 Time: 20:07
Sample: 1981Q3 2000Q1
Included observations: 75
Total system (balanced) observations 150
Residual covariance matrix restricted to be diagonal in FIML estimation
Estimation settings: tol=1.0e-13, derivs=analytic (linear)
Initial Values: C(1)=0.50000, C(2)=0.50000, C(3)=0.50000, C(4)=0.50000, C(5)=0.50000, C(7)=0.50000, C(6)=0.50000
Convergence achieved after 20 iterations
Coefficient covariance computed using observed Hessian

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.397470	0.103096	3.855322	0.0001
C(2)	-0.291735	0.135905	-2.146618	0.0318
C(3)	0.502727	0.110055	4.567963	0.0000
C(4)	-0.268512	0.234349	-1.145780	0.2519
C(5)	0.411424	0.125584	3.276092	0.0011
C(7)	0.074905	0.248953	0.300881	0.7635
C(6)	0.552657	0.408182	1.353946	0.1758

Log likelihood	-165.4936	Schwarz criterion	4.816128
Avg. log likelihood	-1.103291	Hannan-Quinn criter.	4.686195
Akaike info criterion	4.599829		
Determinant residual covariance	0.381504		

Figure 6.7: FIML Estimates (Diagonal Covariance Option) for the Money/GDP Model

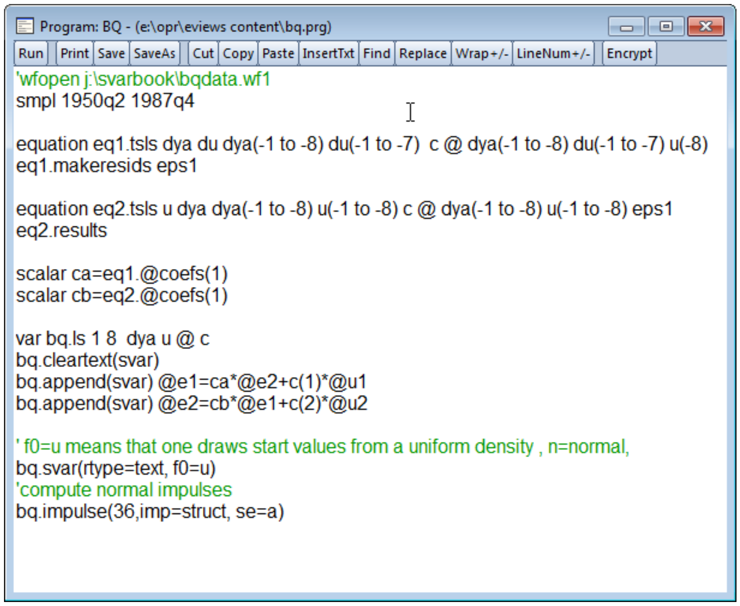


Figure 6.8: EViews Program *bq.prg* to Estimate the Blanchard-Quah Model

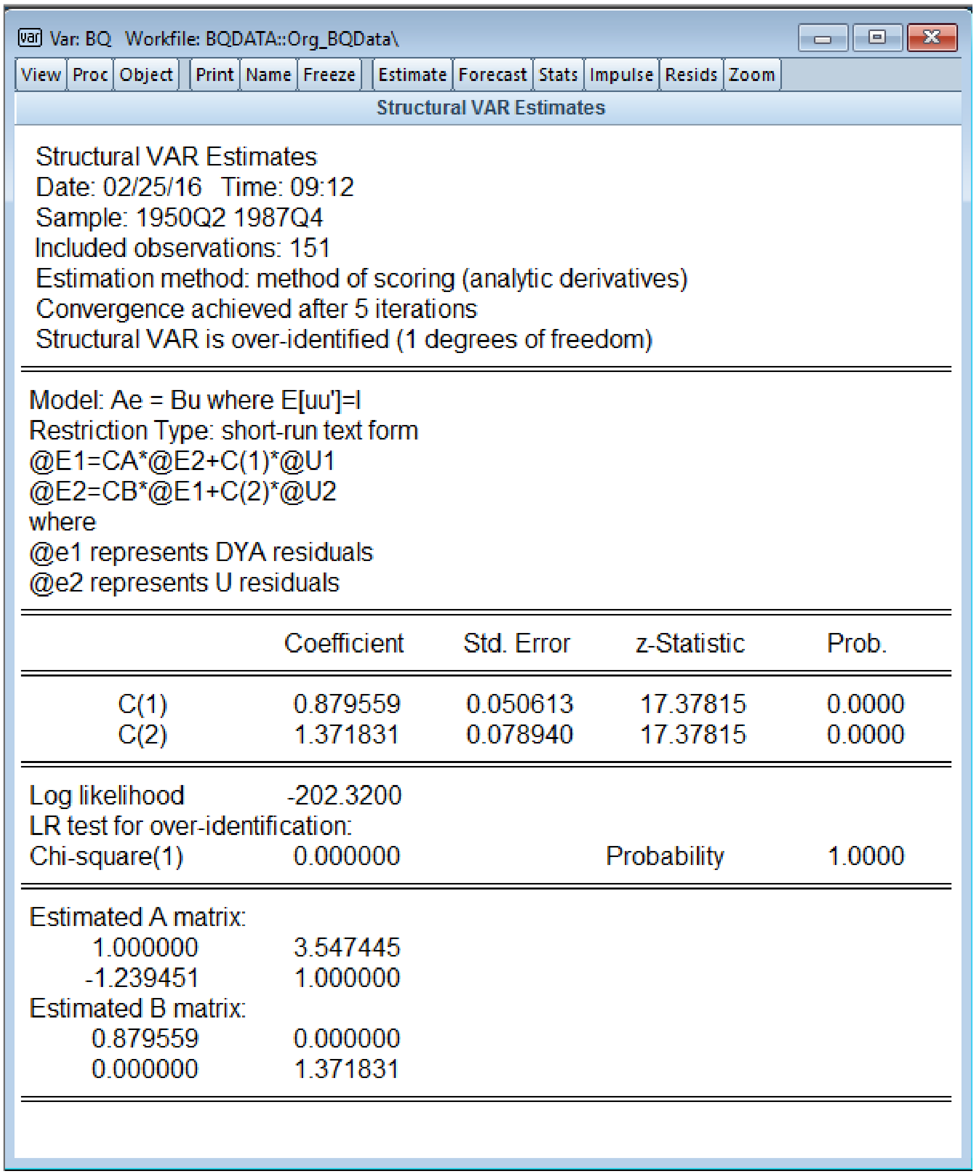
rate. Estimating this system using FIML and a diagonal covariance matrix yields the output shown in Figure 6.12. The sum of the parameter estimates for the lagged coefficients, namely $C(2) + C(4) + C(6) + C(8) + C(10) + C(12) + C(14) + C(16)$, is 3.5474, with a standard error of 1.279038. This matches the estimate for the contemporaneous coefficient on u using the instrumental variable approach and the SVAR routine (see Figure 6.9 and footnote 5).[9]

The two examples we have worked through show that the structural equation with the permanent shock has the other variable in differenced form if the shock in that equation has a zero long-run effect on the first variable. This could occur either because the second shock is permanent with a long-run zero effect on the first variable or it is a transitory shock. This result extends to any number of variables. So, if there is a mixture of $I(1)$ and $I(0)$ variables, and the shocks introduced by the $I(0)$ variables are transitory, then all those variables will appear in differenced form in the equation with the $I(1)$ variables.

6.4.3 Analytical Solution for the Two Variable Case

In the two variable SVAR with long-run restrictions that we have been working with it is useful to get an analytical expression for the estimated a_{12}^0, as this helps later to understand a number of inference issues. We will formally do the

[9] The program *bq_mle.prg* in the MLE sub-directory uses the *optimize*() routine in EViews to implement the FIML estimator for the Blanchard-Quah model by maximizing a user defined likelihood function.



Var: BQ Workfile: BQDATA::Org_BQData\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 02/25/16 Time: 09:12
Sample: 1950Q2 1987Q4
Included observations: 151
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 5 iterations
Structural VAR is over-identified (1 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@E1=CA*@E2+C(1)*@U1
@E2=CB*@E1+C(2)*@U2
where
@e1 represents DYA residuals
@e2 represents U residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.879559	0.050613	17.37815	0.0000
C(2)	1.371831	0.078940	17.37815	0.0000

Log likelihood	-202.3200		
LR test for over-identification:			
Chi-square(1)	0.000000	Probability	1.0000

Estimated A matrix:

1.000000	3.547445
-1.239451	1.000000

Estimated B matrix:

0.879559	0.000000
0.000000	1.371831

Figure 6.9: SVAR/IV Output for the Blanchard-Quah Model

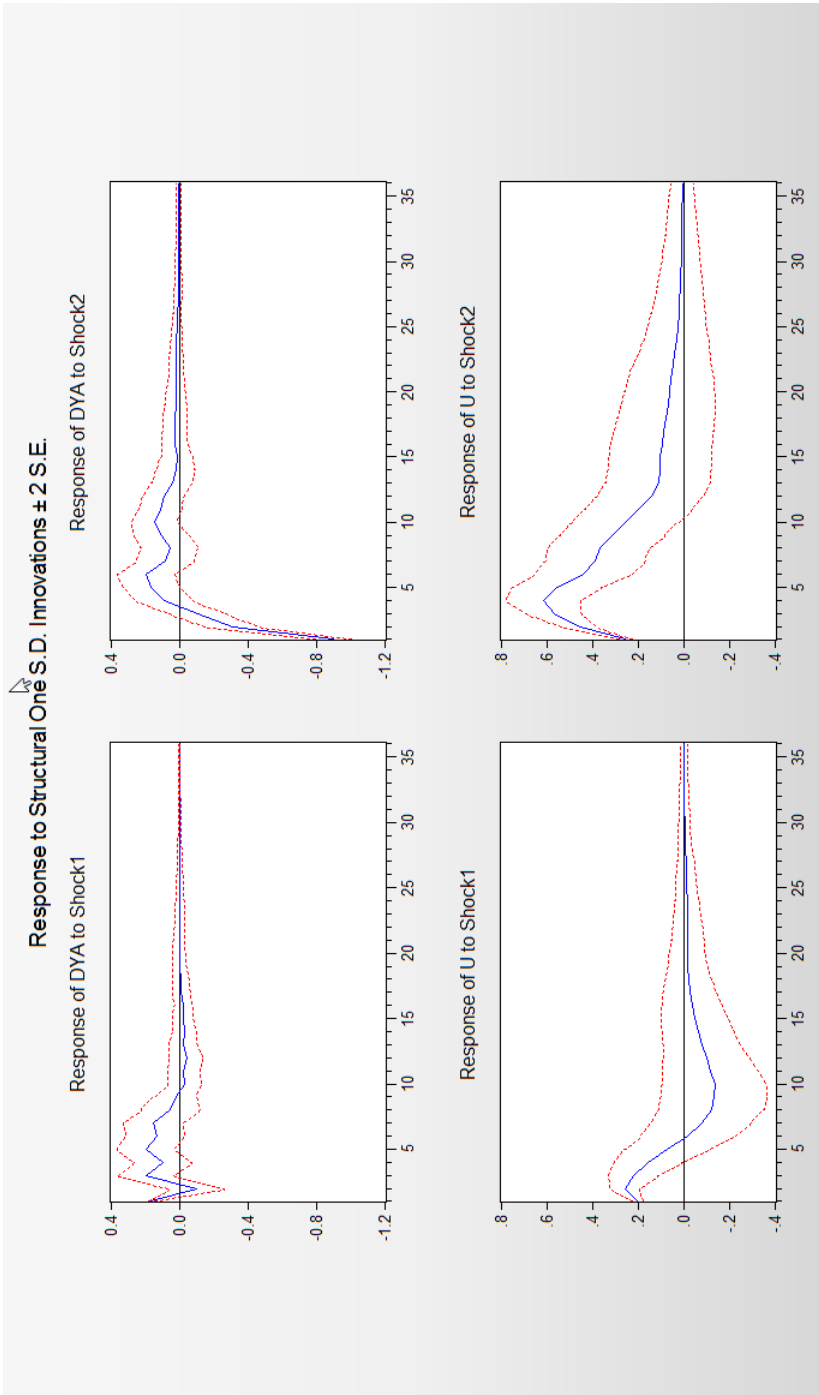
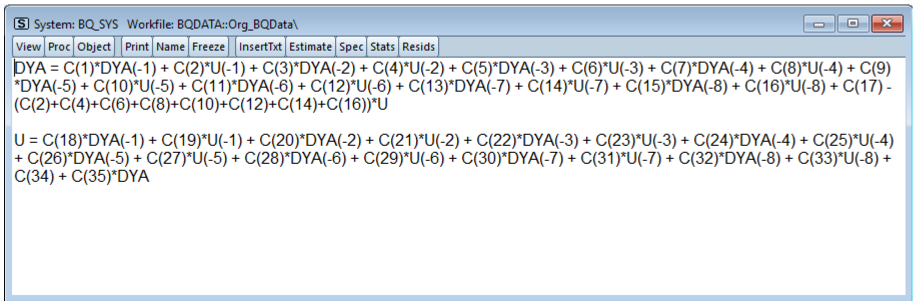


Figure 6.10: Impulse Response Functions for the Blanchard Quah Model




```
DYA = C(1)*DYA(-1) + C(2)*U(-1) + C(3)*DYA(-2) + C(4)*U(-2) + C(5)*DYA(-3) + C(6)*U(-3) + C(7)*DYA(-4) + C(8)*U(-4) + C(9)
*DYA(-5) + C(10)*U(-5) + C(11)*DYA(-6) + C(12)*U(-6) + C(13)*DYA(-7) + C(14)*U(-7) + C(15)*DYA(-8) + C(16)*U(-8) + C(17) -
(C(2)+C(4)+C(6)+C(8)+C(10)+C(12)+C(14)+C(16))*U

U = C(18)*DYA(-1) + C(19)*U(-1) + C(20)*DYA(-2) + C(21)*U(-2) + C(22)*DYA(-3) + C(23)*U(-3) + C(24)*DYA(-4) + C(25)*U(-4)
+ C(26)*DYA(-5) + C(27)*U(-5) + C(28)*DYA(-6) + C(29)*U(-6) + C(30)*DYA(-7) + C(31)*U(-7) + C(32)*DYA(-8) + C(33)*U(-8) +
C(34) + C(35)*DYA
```

Figure 6.11: EViews SYSTEM Object Code to Estimate the Blanchard - Quah Model

$p=1$ case. In this instance $a_{12}^0 + a_{12}^1 = 0$, i.e. $[-A_0 + A_1]_{12} = 0$, where $[F]_{ij}$ means the i,j'th element of F (remember that A_0 is defined as having -a_{ij}^0 $(i \neq j)$ elements) on the off-diagonal and 1 on the diagonal, i.e. $A_0 = \begin{bmatrix} 1 & -a_{12}^0 \\ -a_{21}^0 & 1 \end{bmatrix}$. Now, because $B_j = A_0^{-1} A_j$ implies $A_1 = A_0 B_1$,

$$[-A_0 + A_1]_{12} = [-A_0 + A_0 B_1]_{12}\,.$$

Hence the RHS is the $(1,2)$'th element of the matrix

$$\begin{pmatrix} -1 & a_{12}^0 \\ a_{21}^0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & -a_{12}^0 \\ -a_{21}^0 & 1 \end{pmatrix} \times \begin{pmatrix} b_{11}^1 & b_{12}^1 \\ b_{21}^1 & b_{22}^1 \end{pmatrix}.$$

Putting the (1,2)'th element to zero means

$$a_{12}^0 + b_{12}^1 - a_{12}^0 b_{22}^1 = 0$$

which implies that

$$a_{12}^0 = \frac{-b_{12}^1}{1 - b_{22}^1} = \frac{-[B(1)]_{12}}{[B(1)]_{22}}. \tag{6.6}$$

For an $SVAR(p)$ we would get

$$a_{12}^0 = \frac{-\sum_{j=1}^p b_{12}^j}{1 - \sum_{j=1}^p b_{22}^j}$$

Note that the VAR(1) equation for z_{2t} would be

$$\begin{aligned} z_{2t} &= b_{22}^1 z_{2t-1} + b_{21}^1 z_{1t-1} + e_{2t} \\ \Longrightarrow & \quad \Delta z_{2t} = (b_{22}^1 - 1) z_{2t-1} + b_{21}^1 z_{1t-1} + e_{2t} \\ &= -[B(1)]_{22} z_{2t-1} + b_{21}^1 z_{1t-1} + e_{2t}, \end{aligned}$$

so that the correlation between Δz_{2t} and z_{2t-1} depends on $[B(1)]_{22}$. When this is close to zero, z_{2t-1} would be a *weak instrument* for Δz_{2t} . Consequently, the distribution of the estimator of a_{12}^0, $\hat{a}_{12}^0$ will be affected. It is clear from (6.6) why a small value of $[B(1)]_{22}$ will have a big impact on the density of $\hat{a}_{12}^0$.

System: BQ_SYS Workfile: BQDATA::Org_BQData\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: BQ_SYS
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Date: 04/12/16 Time: 20:15
Sample: 1950Q2 1987Q4
Included observations: 151
Total system (balanced) observations 302
Residual covariance matrix restricted to be diagonal in FIML estimation
Convergence achieved after 104 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.256784	0.253480	-1.013034	0.3110
C(2)	4.150758	2.239455	1.853467	0.0638
C(3)	-0.065862	0.181929	-0.362021	0.7173
C(4)	-0.467450	1.131254	-0.413214	0.6794
C(5)	-0.194283	0.170410	-1.140092	0.2542
C(6)	-0.359713	0.575014	-0.625574	0.5316
C(7)	0.039507	0.123827	0.319050	0.7497
C(8)	-0.029679	0.605371	-0.049027	0.9609
C(9)	0.040176	0.135465	0.296581	0.7668
C(10)	0.132529	0.672223	0.197150	0.8437
C(11)	-0.053082	0.193800	-0.273900	0.7842
C(12)	0.383941	0.650549	0.590180	0.5551
C(13)	-0.090535	0.146497	-0.617995	0.5366
C(14)	-0.412134	0.609838	-0.675809	0.4992
C(15)	0.113277	0.121984	0.928623	0.3531
C(16)	0.149194	0.414051	0.360327	0.7186
C(17)	-0.044998	0.090679	-0.496233	0.6197
C(18)	-0.326684	0.720472	-0.453430	0.6502
C(19)	2.155087	2.683908	0.802966	0.4220
C(20)	-0.347301	0.913503	-0.380186	0.7038
C(21)	-2.701482	6.936652	-0.389450	0.6969
C(22)	-0.004149	0.260370	-0.015935	0.9873
C(23)	1.398498	4.167793	0.335549	0.7372
C(24)	-0.113559	0.394192	-0.288081	0.7733
C(25)	-0.736872	2.180773	-0.337895	0.7354
C(26)	-0.177568	0.531596	-0.334028	0.7384
C(27)	0.397804	1.300591	0.305864	0.7597
C(28)	-0.330406	0.856299	-0.385853	0.6996
C(29)	-0.136276	1.119803	-0.121696	0.9031
C(30)	-0.132171	0.360474	-0.366660	0.7139
C(31)	0.142402	1.158915	0.122875	0.9022
C(32)	0.068896	0.167900	0.410336	0.6816
C(33)	0.233391	0.748477	0.311821	0.7552
C(34)	-0.044572	0.172448	-0.258468	0.7960
C(35)	1.239451	4.144715	0.299044	0.7649

Log likelihood	-184.2845	Schwarz criterion	3.603800
Avg. log likelihood	-0.610214	Hannan-Quinn criter.	3.188551

Figure 6.12: FIML Estimates (Diagonal Covariance Option) for the Blanchard-Quah Model

6.4.4 A Four Variable Model with Permanent Shocks - Peersman (2005)

Peersman (2005) estimates four-variable SVARs for the Euro region and the U.S. The variables are the first difference of the log of oil prices ($z_{1t} = \Delta lpoil_t$), output growth ($z_{2t} = \Delta y_t$), the short term nominal interest rate ($z_{3t} = s_t$) and consumer price inflation ($z_{4t} = \Delta p_t$). The three variables assumed to be $I(1)$ are the log level of oil prices, the log level of output and the log of the price level. As explained earlier these will appear in the SVAR as differences. There is no cointegration between the variables. The short-term nominal interest rate is taken to be $I(0)$. From this description there are at least three permanent shocks along with one extra shock which is associated with the structural equation that is normalized on the $I(0)$ variable s_t. It is necessary to decide on what the effects of this shock are. Peersman treats it as permanent but having zero effects on some variables. In this presentation we deal only with the U.S. data, which involves a quarterly SVAR(3) estimated over the period 1980Q1 to 2002Q2.

There are four shocks in the model and Peersman names these as two supply shocks - the first ε_{1t} being an oil price shock, the second ε_{2t} is labeled a supply shock, ε_{3t} is a monetary policy shock and the fourth ε_{4t} is a demand shock. A combination of short and long run restrictions is used to separate the shocks, with the short run restrictions being:

1. Oil prices are weakly exogenous, i.e. there is no contemporaneous effect of non-oil shocks (ε_{2t}, ε_{3t}, ε_{4t}) upon oil prices z_{1t}. This means that the change in oil prices may be treated as an exogenous regressor in the output, money and inflation equations.

2. Money shocks ε_{3t} have no *contemporaneous* effect on output z_{2t}.

The second set of restrictions are *long-run* in nature. They are

- A permanent demand shock ε_{4t} that has a zero long-run effect on GDP.

- The money shock ε_3 that has a *zero long-run effect on output but a non-zero effect on the other I(1) variables.* It is this assumption that makes it a permanent shock since it would be transitory only if it has a zero long-run effect on *all* the *I(1)* variables.

These assumptions imply that $C(1)$ and the B matrix in EViews have the form

$$C(1) = \begin{bmatrix} * & * & * & * \\ * & * & 0 & 0 \\ * & * & * & * \\ * & * & * & * \end{bmatrix}, \; B = A_0^{-1} = \begin{bmatrix} * & 0 & 0 & 0 \\ * & * & 0 & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix},$$

showing that there are two restrictions on $C(1)$ and four on A_0. The model is therefore exactly identified. The shocks are distinguished in the following way: oil and supply do not have a zero long-run effect on output whereas demand and money shocks do. Oil is differentiated from supply via the exogeneity of the oil price. Money is separated from demand by the fact that the money shock has no contemporaneous effect on output.

Now the VAR underlying Peersman's SVAR has the form of $B(L)z_t = e_t = B\varepsilon_t$ making $C(1) = B^{-1}(1)B = \Psi B$, so that the long-run restrictions which Peersman imposes are $c_{23}(1) = 0$ and $c_{24}(1) = 0$. This implies the following constraints on the elements of the EViews matrix B (ψ^{ij} are elements of Ψ)

$$\psi^{21}(1)b_{13} + \psi^{22}(1)b_{23} + \psi^{23}(1)b_{33} + \psi^{24}(1)b_{43} = 0 \quad (6.7)$$
$$\psi^{21}(1)b_{14} + \psi^{22}(1)b_{24} + \psi^{23}(1)b_{34} + \psi^{24}(1)b_{44} = 0. \quad (6.8)$$

Moreover, there are the short-run constraints specified above which constrain $b_{12} = 0, b_{13} = 0, b_{14} = 0, b_{23} = 0$. Peersman used (6.7) and (6.8) along with the constraints on b to get his estimates.

An alternative approach which imposes all the restrictions is to use the Shapiro/Watson method. Since Peersman's assumptions imply four structural equations we look at each in turn. .

1. *Oil price inflation*

$$\Delta lpoil_t = lags + \varepsilon_{1t},$$

where "lags" are in all variables (including s_t). We can therefore run OLS on the variables to get $\hat{\varepsilon}_{1t}$.

2. *Output growth equation*

$$\Delta y_t = a_{21}^0 \Delta lpoil_t + a_{23}^0 \Delta s_t + a_{24}^0 \Delta(\Delta lp_t) + lags + \varepsilon_{2t}$$

Here $\Delta(\Delta lp_t)$ reflects the assumption of a zero long-run effect of demand shocks on output and s_t appears in differenced form, as money has no permanent effect upon output. One can therefore use $\Delta lp_{t-1}, s_{t-1}, \hat{\varepsilon}_{1t}$ as the instruments to estimate the equation and thereby get $\hat{\varepsilon}_{2t}$.

3. *Interest rate equation*

$$s_t = a_{31}^0 \Delta lpoil_t + a_{32}^0 \Delta y_t + a_{34}^0 \Delta lp_t + lags + \varepsilon_{3t}$$

We have $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ as instruments for $\Delta lpoil_t$ and Δy_t, but need one for Δlp_t. Because it is assumed that monetary shocks have a zero contemporaneous effect on output, the reduced form VAR errors e_{2t} (those for the Δy_t equation in the VAR) are uncorrelated with ε_{3t}, allowing the residuals $\hat{e}_{2t}$ to be used as the extra instrument.[10]

[10]It might be thought that a better instrument would be $\hat{e}_{4t}$, which would come from the assumption that monetary shocks had a zero contemporaneous effect on prices. This would make more sense from a New Keynesian model perspective where the monetary effect on prices follows after the effect on output.

4. *Inflation equation*

$$\Delta lp_t = a_{41}^0 \Delta lpoil_t + a_{42}^0 \Delta y_t + a_{43}^0 s_t + lags + \varepsilon_{4t}$$

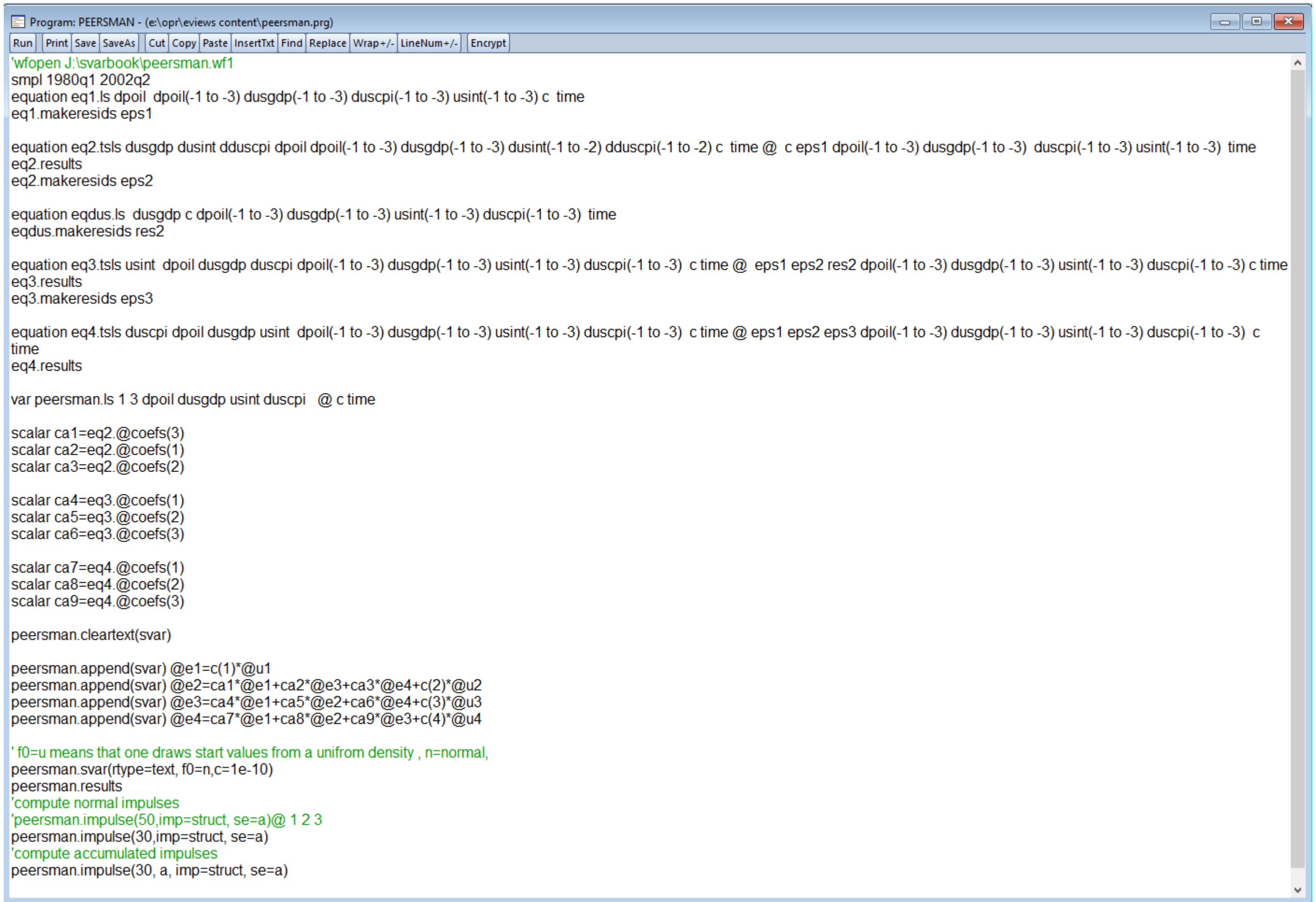
In this final equation all residuals $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ and $\hat{\varepsilon}_{3t}$ will be the instruments to estimate the equation.

The code to do this estimation using the IV approach is in *peersman.prg,* the contents of which are shown in Figure 6.13. The results from the SVAR routine are shown in Figure 6.14, while accumulated impulse responses for the levels of the price of oil, the cpi and output in response to the four shocks are given in Figures 6.15, 6.16, 6.17. These agree with what Peersman shows.

Clearly the interest rate shock has a long-run effect on the price of oil and the price level (CPI), which might be thought undesirable. The values to which the impulse responses converge are well away from zero and are different from one another, so that the real price of oil changes in the long-run in response to a one-period interest rate shock. Certainly we would not expect a nominal shock to change the relative price of oil to the CPI, and in Fisher *et al.* (2015) a restriction is imposed that the monetary shock cannot affect the real oil price in the long-run. It is found that, while there are only minor price puzzles for Peersman's original model, this changes when monetary shocks are restricted to have a zero long-run impact on the real oil price.

Estimating Peersman's model directly using FIML reveals an implicit restriction coming from the identification assumptions adopted by Peersman (2005), one that is automatically enforced by the IV approach. First, with respect to the output growth equation, one needs to constrain the sum of the corresponding lagged coefficients on s_t and Δlp_t to be equal but opposite in sign to the contemporaneous coefficients a_{23}^0 and a_{24}^0. Doing so ensures that $c_{23}(1) = 0$ and $c_{24}(1) = 0$ and, as noted above, provides two of the six identifying restrictions needed to estimate the model. An additional three constraints come from the oil price equation which ensure that oil prices are exogeneous to output, interest rates and inflation. Clearly, an additional constraint on the parameters of the model is needed to achieve exact identification. This is that money shocks, ε_{3t}, do not have a contemporaneous effect on output, z_{2t}, i.e., $[B]_{23} = 0$. It means that the coefficient on the interest rate variable in the inflation equation, a_{43}^0, must be constrained to equal $-\frac{a_{23}^0}{a_{24}^0}$. This requirement can be given an intuitive explanation. Starting from the inflation equation, a shock to the interest rate increases inflation by a_{43}^0 in the current period. In the absence of any constraints on the output growth equation, this would affect output growth by $a_{24}^0 * a_{43}^0$ through the inflation channel (i.e. $\Delta(\Delta lp)_t$). To offset this, the coefficient on the interest rate term in the output growth equation, a_{23}^0, must be constrained to $-a_{24}^0 * a_{43}^0$. Now since a_{23}^0 and a_{24}^0 are already used to constrain the long-run impact of monetary and demand shocks respectively, the only way to enforce the last condition jointly with the constraints on a_{23}^0 and a_{24}^0 is to set $a_{43}^0 = -\frac{a_{23}^0}{a_{24}^0}$.[11]

[11] It can be shown algebriacally that this condition ensures that $[B]_{23} = \left[A_0^{-1}\right]_{23} = 0$.



```
Program: PEERSMAN - (e:\opr\eviews content\peersman.prg)
Run Print Save SaveAs Cut Copy Paste InsertTxt Find Replace Wrap+/- LineNum+/- Encrypt

'wfopen J:\svarbook\peersman.wf1
smpl 1980q1 2002q2
equation eq1.ls dpoil  dpoil(-1 to -3) dusgdp(-1 to -3) duscpi(-1 to -3) usint(-1 to -3) c  time
eq1.makeresids eps1

equation eq2.tsls dusgdp dusint dduscpi dpoil dpoil(-1 to -3) dusgdp(-1 to -3) dusint(-1 to -2) dduscpi(-1 to -2) c  time @  c eps1 dpoil(-1 to -3) dusgdp(-1 to -3)  duscpi(-1 to -3) usint(-1 to -3)  time
eq2.results
eq2.makeresids eps2

equation eqdus.ls  dusgdp c dpoil(-1 to -3) dusgdp(-1 to -3) usint(-1 to -3) duscpi(-1 to -3)  time
eqdus.makeresids res2

equation eq3.tsls usint  dpoil dusgdp duscpi dpoil(-1 to -3) dusgdp(-1 to -3) usint(-1 to -3) duscpi(-1 to -3)  c time @  eps1 eps2 res2 dpoil(-1 to -3) dusgdp(-1 to -3) usint(-1 to -3) duscpi(-1 to -3) c time
eq3.results
eq3.makeresids eps3

equation eq4.tsls duscpi dpoil dusgdp usint  dpoil(-1 to -3) dusgdp(-1 to -3) usint(-1 to -3) duscpi(-1 to -3)  c time @ eps1 eps2 eps3 dpoil(-1 to -3) dusgdp(-1 to -3) usint(-1 to -3) duscpi(-1 to -3)  c
time
eq4.results

var peersman.ls 1 3 dpoil dusgdp usint duscpi   @ c time

scalar ca1=eq2.@coefs(3)
scalar ca2=eq2.@coefs(1)
scalar ca3=eq2.@coefs(2)

scalar ca4=eq3.@coefs(1)
scalar ca5=eq3.@coefs(2)
scalar ca6=eq3.@coefs(3)

scalar ca7=eq4.@coefs(1)
scalar ca8=eq4.@coefs(2)
scalar ca9=eq4.@coefs(3)

peersman.cleartext(svar)

peersman.append(svar) @e1=c(1)*@u1
peersman.append(svar) @e2=ca1*@e1+ca2*@e3+ca3*@e4+c(2)*@u2
peersman.append(svar) @e3=ca4*@e1+ca5*@e2+ca6*@e4+c(3)*@u3
peersman.append(svar) @e4=ca7*@e1+ca8*@e2+ca9*@e3+c(4)*@u4

' f0=u means that one draws start values from a unifrom density , n=normal,
peersman.svar(rtype=text, f0=n,c=1e-10)
peersman.results
'compute normal impulses
'peersman.impulse(50,imp=struct, se=a)@ 1 2 3
peersman.impulse(30,imp=struct, se=a)
'compute accumulated impulses
peersman.impulse(30, a, imp=struct, se=a)
```

Figure 6.13: EViews Program *peersman.prg* to Replicate Peersman (2005)

Structural VAR Estimates
Date: 02/04/16 Time: 06:24
Sample: 1980Q1 2002Q2
Included observations: 90
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 6 iterations
Structural VAR is over-identified (6 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@E1=C(1)*@U1
@E2=CA1*@E1+CA2*@E3+CA3*@E4+C(2)*@U2
@E3=CA4*@E1+CA5*@E2+CA6*@E4+C(3)*@U3
@E4=CA7*@E1+CA8*@E2+CA9*@E3+C(4)*@U4
where
@e1 represents DPOIL residuals
@e2 represents DUSGDP residuals
@e3 represents USINT residuals
@e4 represents DUSCPI residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	14.11005	1.051701	13.41641	0.0000
C(2)	0.647575	0.048267	13.41641	0.0000
C(3)	1.024799	0.076384	13.41641	0.0000
C(4)	0.387080	0.028851	13.41641	0.0000

Log likelihood	-547.3138			
LR test for over-identification:				
Chi-square(6)	7.99E-14		Probability	1.0000

Estimated A matrix:

1.000000	0.000000	0.000000	0.000000
0.015836	1.000000	-0.297656	-1.898772
0.008121	-0.498148	1.000000	-2.701259
-0.009260	0.367328	0.156762	1.000000

Estimated B matrix:

14.11005	0.000000	0.000000	0.000000
0.000000	0.647575	0.000000	0.000000
0.000000	0.000000	1.024799	0.000000
0.000000	0.000000	0.000000	0.387080

Figure 6.14: Structural VAR Estimates for Peersman (2005) Using *peersman.prg*

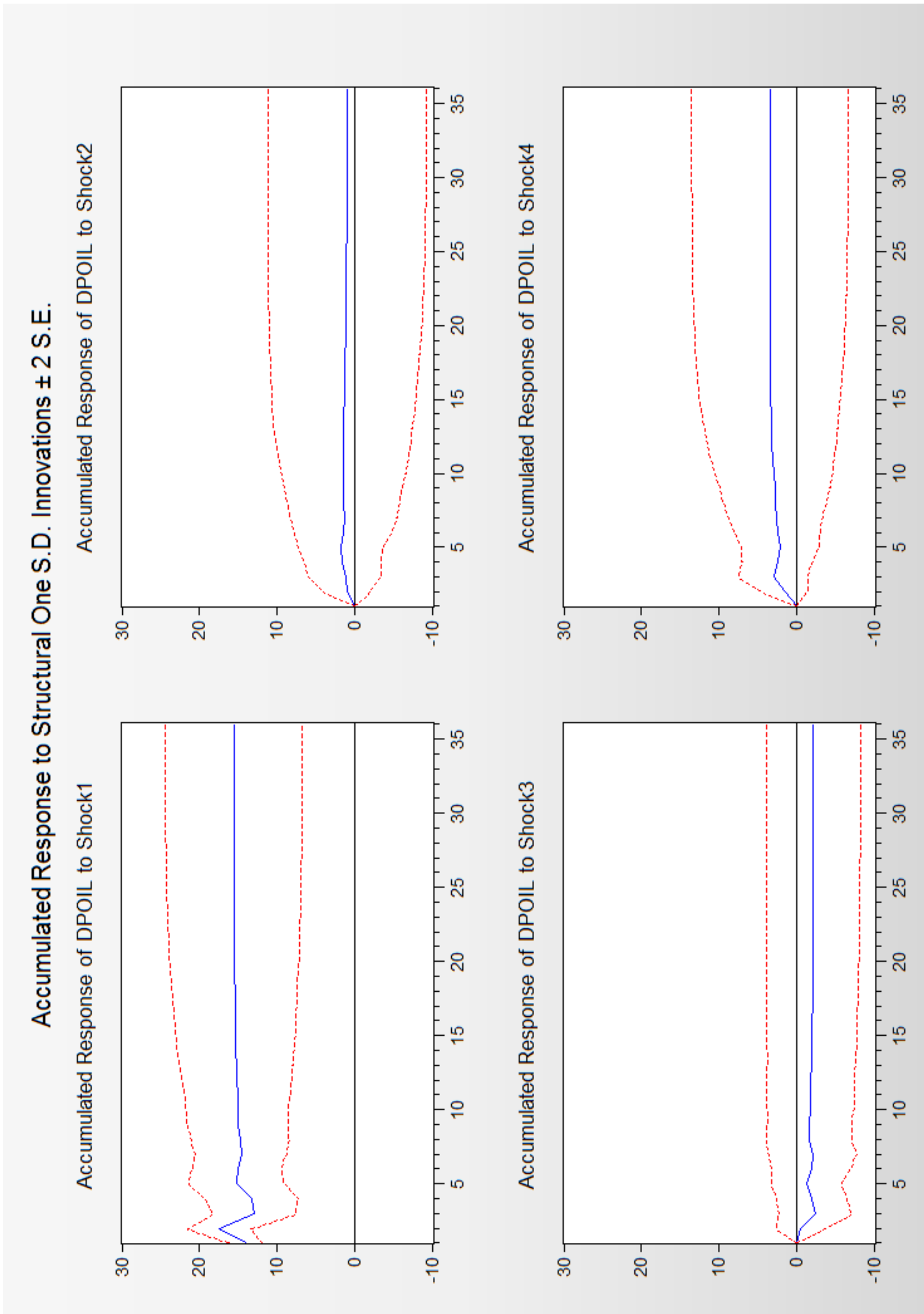


Figure 6.15: Accumulated Impulse Responses of Levels of the Price of Oil: Peersman's(2005) Model

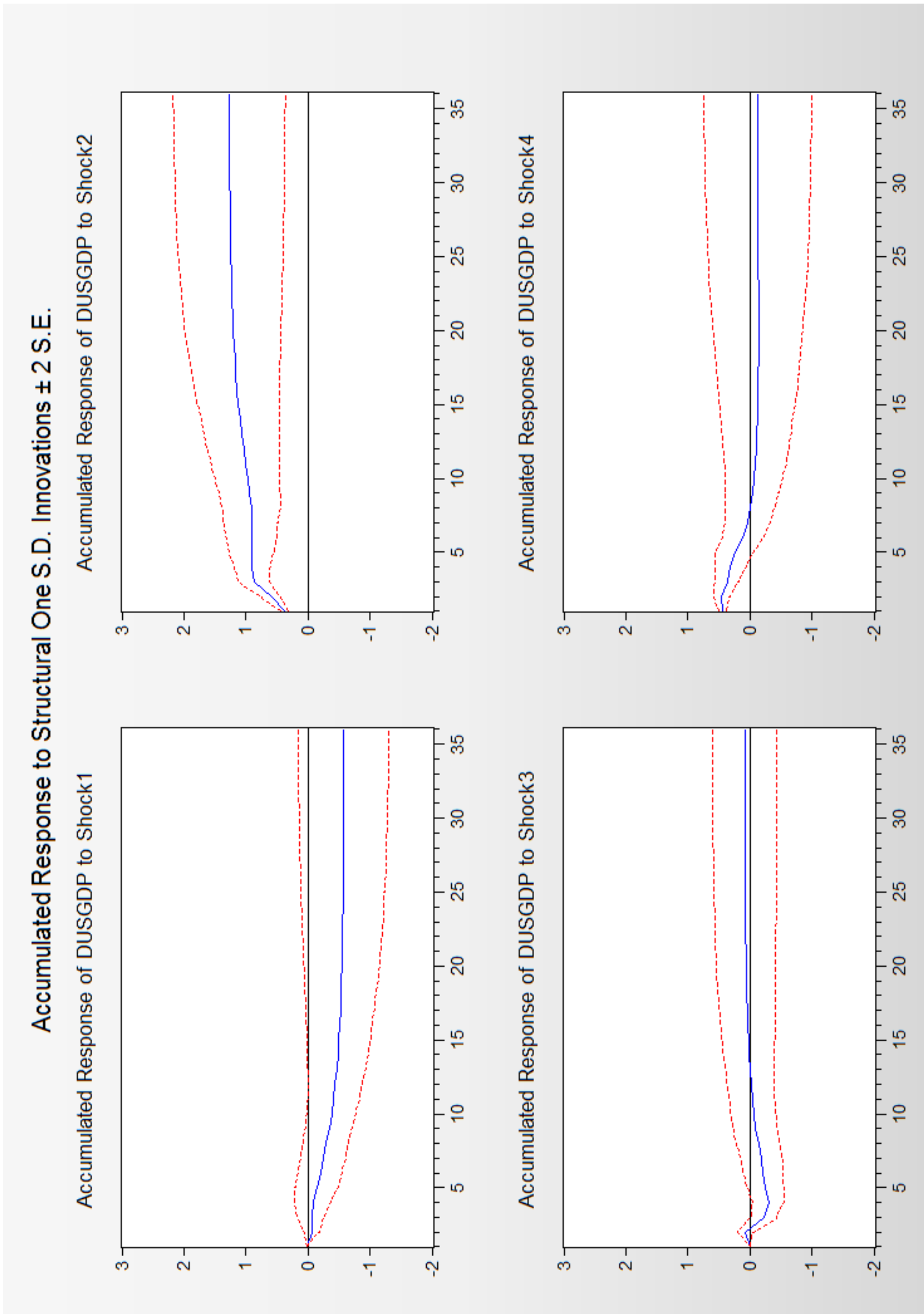


Figure 6.16: Accumulated Impulse Responses of Output in Peersman's (2005) Model

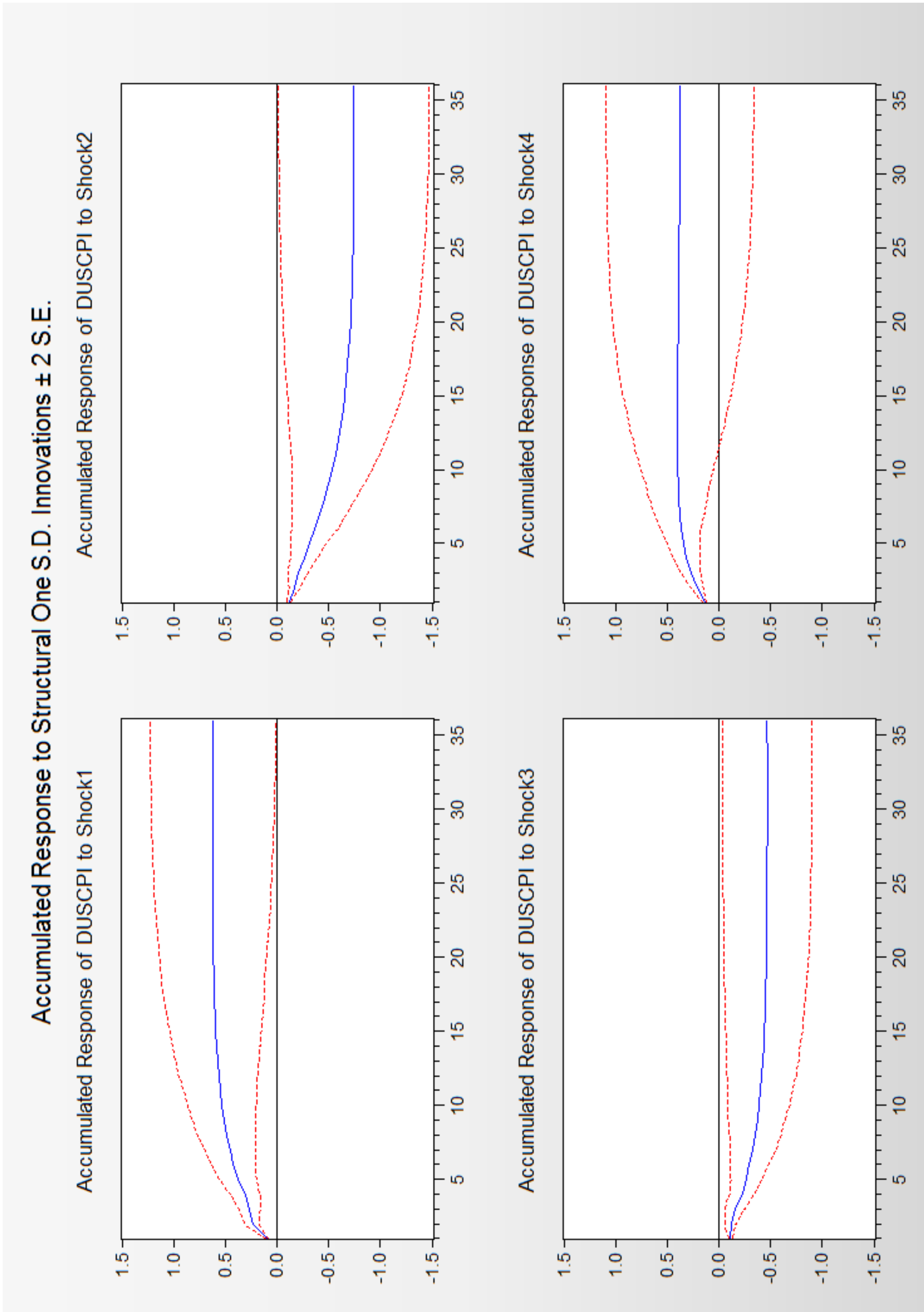


Figure 6.17: Accumulated Impulse Responses of the CPI in Peersman's (2005) Model

The system object code to estimate the Peersman model is given in Figure 6.18. Estimating the model using FIML and a diagonal covariance matrix yields, for example, C(21)=-0.329765, C(22)=-0.205430 and C(23)=0.237539, implying a contemporaneous coefficient estimate for a_{23}^0 of $(C(21) + C(22) + C(23)) = -0.297656$, with a standard error of 0.142151 computed using the delta method. Likewise, the estimate of a_{24} can be obtained from $(C(24)+C(25)+C(26))$. This gives $a_{24} = -1.898772$ with a standard error of 0.830540. Lastly, note that $\frac{a_{23}^0}{a_{24}^0} = 0.156762$. All these estimates match those obtained using the instrumental variable/SVAR approach.[12]

This model was chosen to illustrate the point that working with $I(1)$ variables in difference form (which is appropriate), while having an $I(0)$ variable in levels in the SVAR, will lead to the shock connected to the $I(0)$ variable having a permanent effect on the level of the $I(1)$ variables, unless steps are taken to ensure that this does not happen. Thus if we wanted a long-run zero response of the price level to the monetary shock then the fourth equation above would have to have Δs_t as a regressor rather than s_t. There seem to be many SVAR studies with this difficulty; some of these are mentioned in Fisher *et al.* (2015).

6.4.5 Revisiting the Small Macro Model with a Permanent Supply Shock

The small macro model had issues with price puzzles. It is interesting to see what happens if it is assumed that the output gap measured by Cho and Moreno (which took a deterministic trend out of the log of GDP) is $I(1)$ rather than $I(0)$. The other two variables are still treated as being I(0).[13]

This will mean that one permanent supply side shock is present in the system and we will assume that there are two transitory shocks. Then the SVAR will consist of Δy_t (the change in the output gap), the interest rate (i_t) and inflation (π_t). Here y_t is $I(1)$ and π_t, i_t are $I(0)$. Assuming an SVAR(2), the specification of the system would be

$$
\begin{aligned}
\Delta y_t &= a_{12}^0 \Delta i_t + a_{13}^0 \Delta \pi_t + a_{11}^1 \Delta y_{t-1} + a_{12}^1 \Delta i_{t-1} + a_{13}^1 \Delta \pi_{t-1} \\
&\quad a_{11}^2 \Delta y_{t-2} + \varepsilon_{1t} \qquad (6.9) \\
i_t &= a_{21}^0 \Delta y_t + a_{23}^0 \pi_t + a_{21}^1 \Delta y_{t-1} + a_{22}^1 i_{t-1} + a_{23}^1 \pi_{t-1} \\
&\quad a_{21}^2 \Delta y_{t-2} + a_{22}^2 i_{t-2} + a_{23}^2 \pi_{t-2} + \varepsilon_{2t} \qquad (6.10) \\
\pi_t &= a_{31}^0 \Delta y_t + a_{32}^0 i_t + a_{31}^1 \Delta y_{t-1} + a_{33}^1 \pi_{t-1} + a_{32}^1 i_{t-1} + \\
&\quad a_{31}^2 \Delta y_{t-2} + a_{32}^2 i_{t-2} + a_{33}^2 \pi_{t-2} + \varepsilon_{3t}, \qquad (6.11)
\end{aligned}
$$

where Δi_t and $\Delta \pi_t$ in the first equation ensure that the second and third shocks are transitory.

[12]See also the program code in *peersman_mle.prg* in the MLE sub-directory for an equivalent approach that uses *optimize*() and a user-defined likelihood function.

[13]This means that the underlying process for GDP growth is $\Delta y_t = b + v_t$ where v_t is I(0).

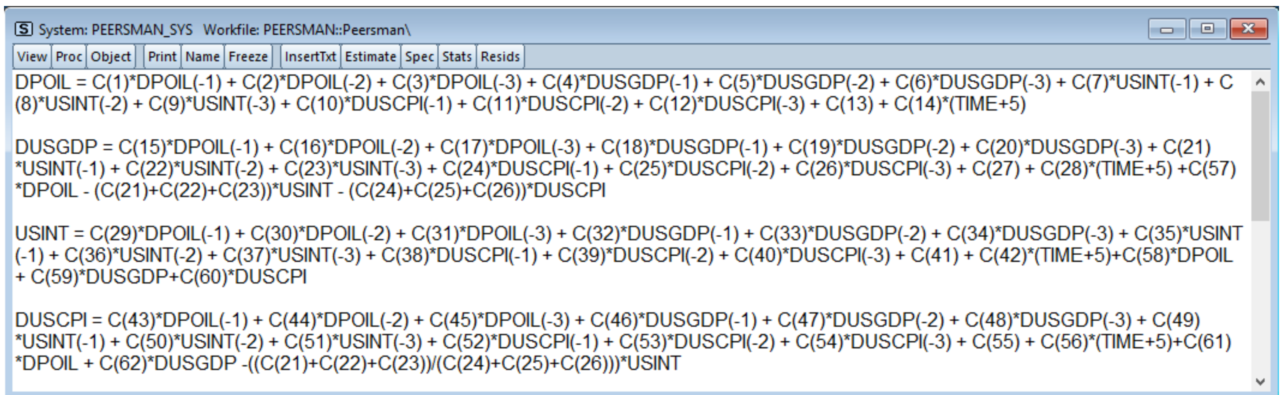


Figure 6.18: System Object Code to Estimate Peersman (2005)

Equation 6.9 can be estimated using π_{t-2}, i_{t-2} (as well as the lagged values) as instruments, producing residuals $\hat{\varepsilon}_{1t}$. To estimate the interest rate equation (6.10) we have lagged values of the variables and $\hat{\varepsilon}_{1t}$, but this leaves us one instrument short. In the recursive system of Chapter 4 interest rates had a zero contemporaneous effect on output, so we impose that again. However we do not make this system recursive. Rather we exploit the fact established in Chapter 4 that, under such an assumption about monetary effects, the residuals from the VAR equation for Δy_t, namely $\hat{e}_{1t}$, can be used as an instrument in the interest rate equation. Therefore, we estimate that equation using $\hat{\varepsilon}_{1t}$, $\hat{e}_{1t}$ and the lagged values of variables as instruments. This gives the residuals $\hat{\varepsilon}_{2t}$ which can be used along with $\hat{\varepsilon}_{1t}$ to estimate the inflation equation. *Chomoreno_perm.prg* in Figure 6.19 provides the code to do this.

The IV/SVAR results are shown in Figure 6.20 and the corresponding impulse responses are presented in Figure 6.21. It is interesting to note that, with the exception of a small rise in output to a positive interest rate shock, the results are much closer to what we would have expected than what was found with the recursive model. That model treated all the data as $I(0)$, showing the importance of getting a correct specification of the system represented by the SVAR.

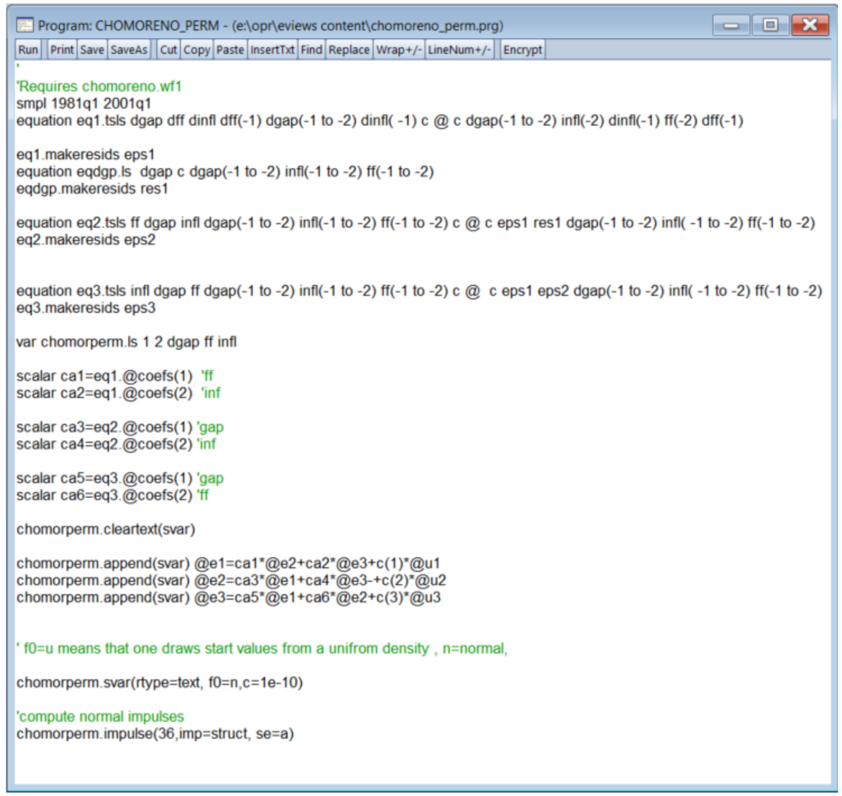
One may also estimate the system using FIML. The required system object code is shown in Figure 6.22. As before, three restrictions are needed to exactly identify the system. Two come from the long-run constraints, which are enforced by restricting a_{12} (the contemporaneous coefficient for inflation) to be equal to the negative of the sum of the lagged coefficients on inflation, namely $-(C(3)+C(4))$, and a_{13} to be equal to $-(C(5)+C(6))$. The requirement that interest rates have a zero contemporaneous effect on output implies that the coefficient on the interest rate in the inflation equation, namely a_{32}^0, must equal $-\left(\frac{a_{12}^0}{a_{13}^0}\right)$. This condition ensures that the contemporaneous interest rate effect on output coming from the inflation channel will be $-a_{12}^0$, which will exactly offset the contemporaneous interest rate effect on output from the (more direct) interest rate channel itself, namely a_{12}^0.

Estimating the system using FIML and a diagonal covariance matrix yields the results shown in Figure 6.23. The estimates are identical to the IV/SVAR results presented in Figure 6.20. For example, the implied coefficient estimate on the interest rate in the output growth equation is $-(C(5)+C(6)) = 0.369652$, and that for the interest rate in the inflation equation is 0.84695.[14]

6.5 Problems with Measuring Uncertainty in Impulse Responses

Can long-run restrictions produce estimators of the quantities of interest that are reliable? Two issues arise in answering this question.

[14] See *chomoremo_perm_mle.prg* in the MLE sub-directory for an equivalent approach that uses *optimize*() and a user-defined likelihood function.



Program: CHOMORENO_PERM - (e:\opr\eviews content\chomoreno_perm.prg)

Run Print Save SaveAs Cut Copy Paste InsertTxt Find Replace Wrap+/- LineNum+/- Encrypt

```
'
'Requires chomoreno.wf1
smpl 1981q1 2001q1
equation eq1.tsls dgap dff dinfl dff(-1) dgap(-1 to -2) dinfl( -1) c @ c dgap(-1 to -2) infl(-2) dinfl(-1) ff(-2) dff(-1)

eq1.makeresids eps1
equation eqdgp.ls  dgap c dgap(-1 to -2) infl(-1 to -2) ff(-1 to -2)
eqdgp.makeresids res1

equation eq2.tsls ff dgap infl dgap(-1 to -2) infl(-1 to -2) ff(-1 to -2) c @ c eps1 res1 dgap(-1 to -2) infl( -1 to -2) ff(-1 to -2)
eq2.makeresids eps2


equation eq3.tsls infl dgap ff dgap(-1 to -2) infl(-1 to -2) ff(-1 to -2) c @  c eps1 eps2 dgap(-1 to -2) infl( -1 to -2) ff(-1 to -2)
eq3.makeresids eps3

var chomorperm.ls 1 2 dgap ff infl

scalar ca1=eq1.@coefs(1)  'ff
scalar ca2=eq1.@coefs(2)  'inf

scalar ca3=eq2.@coefs(1) 'gap
scalar ca4=eq2.@coefs(2) 'inf

scalar ca5=eq3.@coefs(1) 'gap
scalar ca6=eq3.@coefs(2) 'ff

chomorperm.cleartext(svar)

chomorperm.append(svar) @e1=ca1*@e2+ca2*@e3+c(1)*@u1
chomorperm.append(svar) @e2=ca3*@e1+ca4*@e3-+c(2)*@u2
chomorperm.append(svar) @e3=ca5*@e1+ca6*@e2+c(3)*@u3


' f0=u means that one draws start values from a unifrom density , n=normal,

chomorperm.svar(rtype=text, f0=n,c=1e-10)

'compute normal impulses
chomorperm.impulse(36,imp=struct, se=a)
```

Figure 6.19: EViews program *chomoreno_perm.prg* to Allow for One Permanent and Two Transitory Shocks in the Small Macro Model

Structural VAR Estimates
Date: 02/04/16 Time: 07:05
Sample: 1981Q4 2000Q1
Included observations: 74
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 5 iterations
Structural VAR is over-identified (3 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@E1=CA1*@E2+CA2*@E3+C(1)*@U1
@E2=CA3*@E1+CA4*@E3-+C(2)*@U2
@E3=CA5*@E1+CA6*@E2+C(3)*@U3
where
@e1 represents DGAP residuals
@e2 represents FF residuals
@e3 represents INFL residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.726430	0.059712	12.16553	0.0000
C(2)	0.734593	0.060383	12.16553	0.0000
C(3)	1.143837	0.094023	12.16553	0.0000

Log likelihood	-215.2684		
LR test for over-identification:			
Chi-square(3)	3.29E-14	Probability	1.0000

Estimated A matrix:

1.000000	-0.369652	-0.436447
-0.484235	1.000000	-0.665990
1.150318	0.846957	1.000000

Estimated B matrix:

0.726430	0.000000	0.000000
0.000000	-0.734593	0.000000
0.000000	0.000000	1.143837

Figure 6.20: Structural VAR Estimates of the Small Macro Model With One Permanent Shock and Two Transitory Shocks

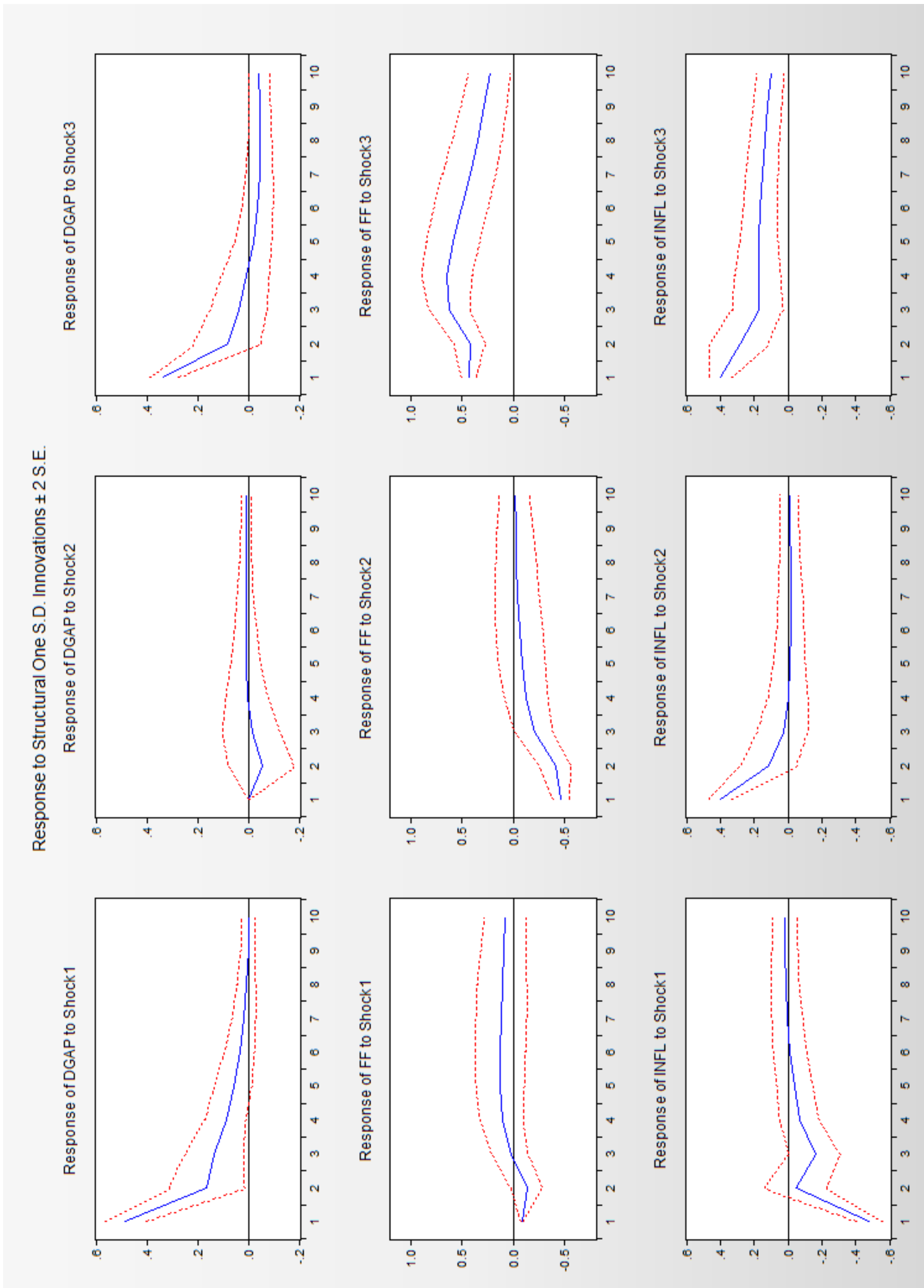


Figure 6.21: Impulse Responses for the Small Macro Model with One Permanent and Two Transitory Shocks

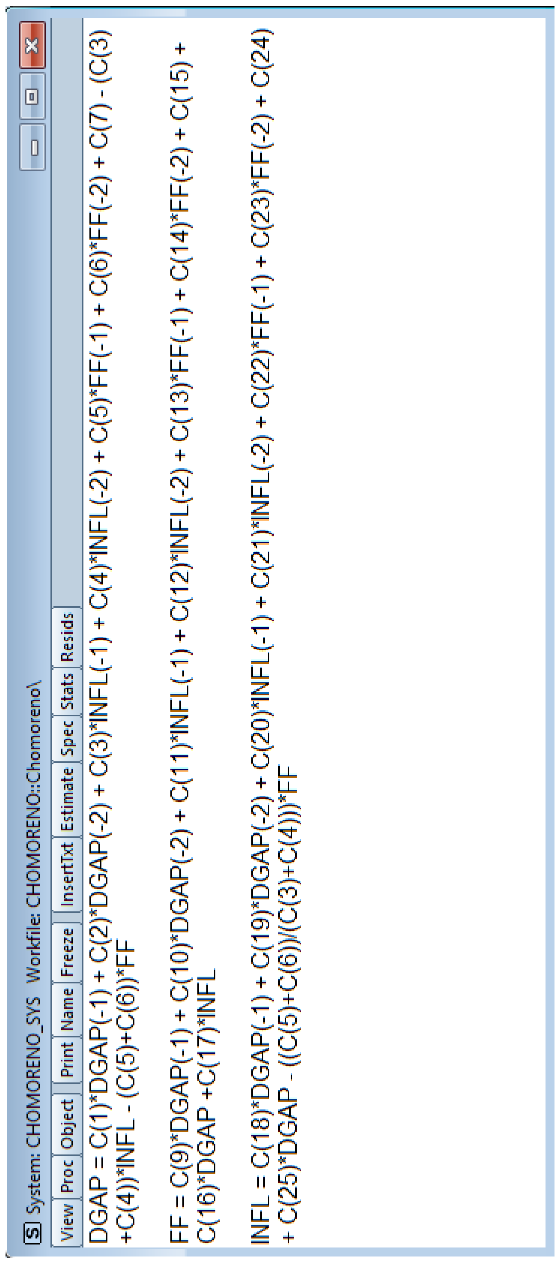


Figure 6.22: SYSTEM Object Code to Estimate the Small Macro Model with One Permanent and Two Transitory Shocks

System: CHOMORENO_SYS Workfile: CHOMORENO::Chomoreno\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: CHOMORENO_SYS
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Date: 04/12/16 Time: 20:24
Sample: 1981Q4 2000Q1
Included observations: 74
Total system (balanced) observations 222
Residual covariance matrix restricted to be diagonal in FIML estimation
Convergence achieved after 117 iterations
Coefficient covariance computed using observed Hessian

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.158337	0.185477	0.853674	0.3933
C(2)	0.047650	0.191235	0.249171	0.8032
C(3)	-0.265014	0.186009	-1.424738	0.1542
C(4)	-0.171433	0.144227	-1.188629	0.2346
C(5)	-0.337427	0.301856	-1.117841	0.2636
C(6)	-0.032224	0.118086	-0.272888	0.7849
C(7)	0.114214	0.094548	1.208008	0.2270
C(9)	-0.375210	0.218265	-1.719055	0.0856
C(10)	0.622947	0.157929	3.944483	0.0001
C(11)	-0.148312	0.185964	-0.797531	0.4251
C(12)	0.153594	0.132227	1.161589	0.2454
C(13)	0.867812	0.129006	6.726892	0.0000
C(14)	-0.161784	0.113688	-1.423047	0.1547
C(15)	-0.071770	0.098924	-0.725502	0.4681
C(16)	0.484235	0.188271	2.572011	0.0101
C(17)	0.665990	0.409320	1.627065	0.1037
C(18)	0.554275	0.287459	1.928191	0.0538
C(19)	0.449994	0.445639	1.009772	0.3126
C(20)	0.346700	0.188178	1.842397	0.0654
C(21)	0.340529	0.239768	1.420245	0.1555
C(22)	0.947369	0.636473	1.488467	0.1366
C(23)	-0.181561	0.214793	-0.845283	0.3980
C(24)	-0.250043	0.163771	-1.526785	0.1268
C(25)	-1.150318	0.620770	-1.853050	0.0639

Log likelihood	-204.2380	Schwarz criterion	6.915860
Avg. log likelihood	-0.919991	Hannan-Quinn criter.	6.466689
Akaike info criterion	6.168596		
Determinant residual covariance	0.276528		

Figure 6.23: FIML Estimates (Diagonal Covariance Matrix) for Small Macro Model with One Permanent and Two Transitory Shocks

1. Is there a bias in estimators caused by the model specification?
2. Are any bias and inference problems due to weak instruments?

Faust and Leeper (1997) looked at the first question by examining the estimation of $\bar{A} = A_0^{-1}$ rather than A_0 in the case of Blanchard and Quah (1989). Because there is a VAR representation $z_t = B(L)e_t = B(L)A_0^{-1}\varepsilon_t = C(L)\varepsilon_t$ we get $C(L) = B(L)\bar{A}$ and so $C(1) = B(1)\bar{A}$. Hence the long-run restriction $[C(1)]_{12} = \left[B(1)\bar{A}\right]_{12} = 0$ and it is clear that the solution for $\bar{A}$ depends on $B(1)$. Because there are only two unknown coefficients in A_0 (since the diagonal terms are unity) it is possible to fix two of the elements of $\bar{A}$, and it is convenient to set the diagonal elements to unity. Then

$$B(1)\bar{A} = \begin{bmatrix} b_{11}(1) & b_{12}(1) \\ b_{21}(1) & b_{22}(1) \end{bmatrix} \begin{bmatrix} 1 & \bar{a}_{12} \\ \bar{a}_{21} & 1 \end{bmatrix}$$

and so $\left[B(1)\bar{A}\right]_{12} = 0$ means

$$\bar{a}_{12} = \frac{-b_{12}^1(1)}{b_{11}(1)}. \tag{6.12}$$

The dependence of $\bar{a}_{12}$ on the elements of $B(1)$ is then apparent. Because this relation does not depend upon the order of the SVAR chosen it is clear that, if one is to estimate $\bar{a}_{12}$ correctly, it must be the case that the chosen SVAR has an underlying VAR that provides an accurate estimator of the true $B(1)$. Because a finite order SVAR is chosen in empirical work it is possible that this fails to capture any higher order VAR that describes the data. This is clearly an approximation error for the SVAR. It will produce biased estimators of the elements of $\bar{A}$ and, hence, impulse responses. Faust and Leeper (1997) proposed this argument - it was due originally to Sims in his causality work. There have been some suggestions about how one might improve the estimator of $B(1)$ using the fact that $B(1)$ is related to the multivariate spectral density of the series at zero - see Christiano *et al.* (2006).

A second issue arises since, as observed earlier, the estimators of A_0 are essentially IV, and it is possible that one may have weak instruments, which can cause the densities of the estimated coefficients in A_0 to depart substantially from normality in a finite sample.[15] We have already seen an illustration of this in Chapter 4. Moreover, there are good reasons for thinking that this could be a major issue when long-run restrictions are invoked. In the canonical two variable case analyzed by Blanchard and Quah z_{2t-8} is being used as an instrument for Δz_{2t}, and so the correlation will be very weak when z_{2t} is a near integrated process. To assess the quality of this instrument we need to regress du_t (Δz_{2t}) against $\{du_{t-j}\}_{j=1}^{7}$ and u_{t-8}. Then the F statistic that the coefficient on u_{t-8} is zero is much less than 10, which suggests a weak instrument. Figure

[15] Since the asymptotic standard errors for impulse responses reported in EViews assumes normality then this means they must be treated with some caution if there are weak instruments. We noted this in Chapter 4.

6.24 shows that the distributions of $\hat{a}^0_{ij}$ and the impulse responses are non-normal, while Figure 6.25 shows that the problem is that $[B(1)]_{12}$ is very close to zero.[16] Because the weak correlation arises when $[B(1)]_{12}$ is close to zero it is useful to look at the literature that examines the distribution of $\hat{a}^0_{12}$ in the "local to zero" (of $[B(1)]_{12}$) context. There has been some work on this - Gospodinov *et al.* (2010) - but rarely on the impulse responses and outside the bivariate context. A recent exception is Chevillon *et al.* (2015) who combine the Anderson-Rubin (1949) test known to work well with weak instruments with a method of adjusting for the fact that the instruments being used are close to being non-stationary. They look at the Blanchard-Quah model, but also the IS-LM structure estimated by Gali and discussed in the next chapter. The method looks promising.

6.6 Sign Restrictions when there are Permanent and Transitory Shocks

If all variables are integrated and there is no cointegration then all shocks are permanent. In this case the SVAR is in differenced variables and the estimation of the SVAR using sign restrictions is done in the normal way. However if $I(0)$ variables are present in the SVAR some shocks may be transitory, and we need to ask how a mixture of $I(1)$ and $I(0)$ variables changes the two methods for imposing sign restriction information presented in the previous chapter, i.e. the *SRR* and *SRC* methodologies.

In the case of *SRR* one needs to be careful in re-combining impulse responses when they are mixed. Suppose there are permanent basis shocks, η^P_t, and transitory ones, η^T_t. Then in the standard *SRR* approach all η_t will be combined together to produce new shocks η^*_t. But this would involve combining both η^P_t and η^T_t together and the resulting η^*_t must be permanent. The only way to ensure that some of the η^*_t are transitory is if we construct them by just re-combining the η^T_t. This suggests that we form $\hat{\eta}^{P*}_t = Q_P\eta^P_t$, $\hat{\eta}^{T*}_t = Q_T\eta^T_t$, with each Q_P, Q_T coming from Givens or Householder transformations. The problem then comes down to producing base shocks that are uncorrelated and which have the correct number of both permanent and transitory shocks.

The alternative is to use the *SRC* method. This also requires that an SVAR be set up that produces the right number of permanent and transitory shocks, but after doing that it is simply a matter of generating any unknown coefficients. An application follows to show how the system would be set up. This utilizes the small macro model and involves a long-run parametric restriction. It corresponds to the model of Section 6.4.5. However, the short run restriction used there that monetary policy had no contemporaneous effect on output – which served to differentiate the demand and monetary shocks – is replaced by sign restrictions. We compare *SRC* and *SRR* in this case and find that

[16] Graphs show the density of $\hat{a}^0_{ij}$ in the Blanchard-Quah model. Note that the labeling uses $B0$ (rather than $A0$) to represent the contemporaneous coefficient matrix.

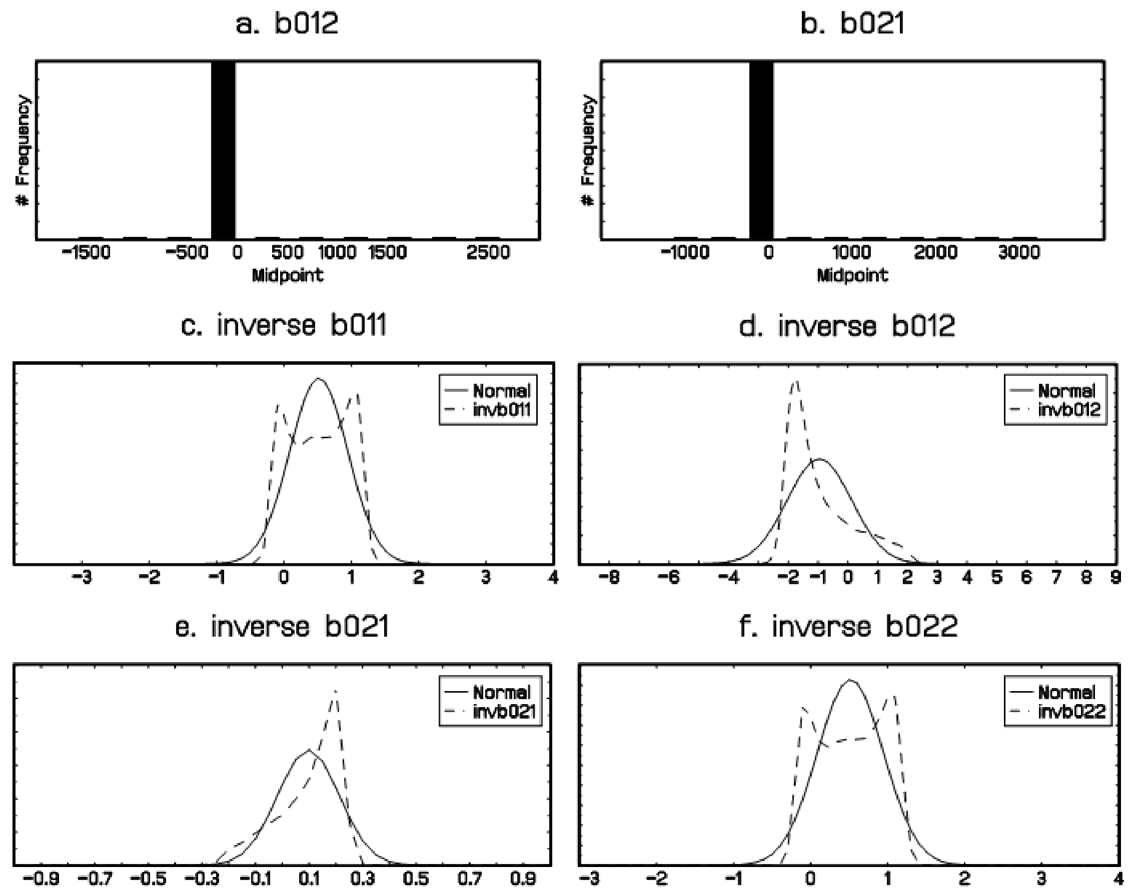


Figure 6.24: Distribution of Structural Parameters and Impact Impulse Responses for the Blanchard-Quah Model

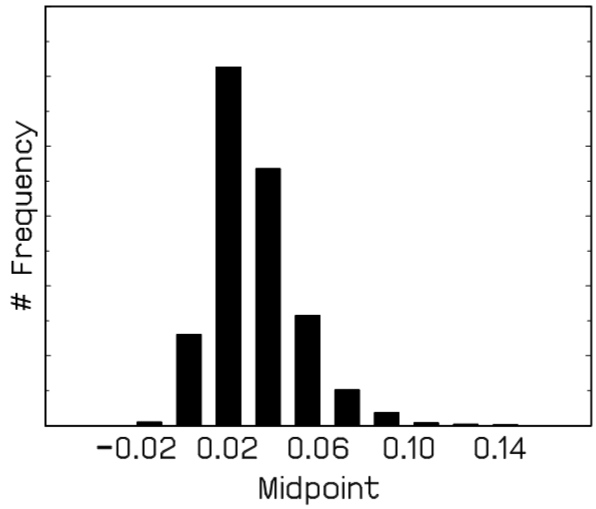


Figure 6.25: Distribution of $[B(1)]_{12}$ in Blanchard and Quah Model

the methodologies produce much the same results. It will become apparent that *SRC* adapts very well to the situation where there are combinations of parametric and sign restrictions.

SRR will find some base transitory shocks in the following way. Set $a^0_{23} = 0$ and then estimate (6.10) using $\hat{\varepsilon}_{1t}$ as an instrument for Δz_{1t}. Then $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ can be used to estimate (6.11) and the shock $\hat{\varepsilon}_{3t}$ follows. The impulse responses for $\hat{\varepsilon}_{2t}$ and $\hat{\varepsilon}_{3t}$ are then recombined to find new transitory shocks. So it is necessary to impose the long run restriction and a recursive assumption to find the initial base shocks.

Now look at the *SRC* methodology. This requires that the second equation (6.10) be estimated and $\hat{\varepsilon}_{1t}, y_{2t-1}, y_{3t-1}$, and Δz_{1t-1} are available as the instruments for this purpose. But this is one fewer instrument than is needed. In Section 6.4.5 that instrument came from the assumption that monetary policy had no contemporaneous effect on GDP. Now that all that we have available are sign restrictions, making it necessary to fix a^0_{23} and create a new dependent variable $y_{2t} - a^0_{23}y_{3t}$. There are now the correct number of instruments and, once the equation is estimated, residuals $\hat{\varepsilon}_{2t}$ would be available. These can be used along with $\hat{\varepsilon}_{1t}, y_{2t-1}, y_{3t-1}$ and Δz_{1t-1} to estimate the last equation. Thus the *SRC* method replaces a^0_{23} with some value, and this is exactly the same situation as occurred with the market model, i.e. once a^0_{23} is replaced by some function of θ, every θ produces a new set of impulse responses . It is crucial to note however that, as θ is varied, the long-run restriction is always enforced by *design of the SVAR,* i.e. by using (6.9) as part of it. Because this parametric (long-run) restriction reduced the number of parameters to be estimated by one, only one parameter needs to be prescribed in order to get all the impulse responses. Sign restrictions are applied to determine which of the two transitory shocks is demand and which is monetary policy. Because the permanent shock does not depend in any way upon the values assigned to a_{23}^0, it is invariant to the changing values of this coefficient, and so it remains the same (just as the SRR impulse responses were invariant to λ). Estimating the SVAR with a permanent shock by the *SRC* technique now results in 45% of the responses satisfying all the sign restrictions, as compared to the 5% with purely transitory shocks.

Chapter 7

SVARs with Cointegrated and I(0) Variables

7.1 Introduction

The previous chapter analyzed the implications for modeling when the variables in a system were integrated and there were permanent and transitory shocks. However, it was assumed that there was no cointegration between the variables. Therefore the I(1) variables in the system appeared in the VAR in first difference form while the I(0) variables were kept in levels. This chapter details how the analysis changes when there is cointegration amongst the I(1) variables. A first difference is that the summative model is now that of a Vector Error Correction Model (VECM) and the interpretative model will be a structural VECM (SVECM). These forms are laid out in the next section. Section 3 then shows how the SVECM can be converted to a SVAR so that the methods introduced in Chapter 6 may be applied. Section 4 works through two examples - Gali's (1999) paper about the impact of productivity shocks and Gali's (1994) IS-LM model. Essentially the focus of this chapter is upon how the isolation of permanent and transitory shocks needs to be performed in a SVECM. Once the method for doing this is understood then the analysis proceeds in the same way as in the previous chapter. Consequently, there are no new issues raised by the imposition of sign restrictions as a way of discriminating between the shocks.

7.2 The VECM and Structural VECM Models

When variables are co-integrated the appropriate summative model will not be the VAR but rather the Vector Error Correction Model (VECM). When there are $r < n$ cointegrating relations in this system the VECM is

$$\Delta\zeta_t = \alpha\beta'\zeta_{t-1} + \Phi_1\Delta\zeta_{t-1} + \mathbf{e}_t, \tag{7.1}$$

where α and β are $n \times r$ matrices (α being the loading matrix and β the co-integrating vectors), $\xi_t = \beta'\zeta_{t-1}$ are the error correction terms, and ζ_t are the (generally log) levels of the $I(1)$ variables. Our strategy will be to transform the information contained in the VECM into a VAR so that the tools discussed in earlier chapters can be utilized. Gali (1992, 1999) seems to have been one of the first to do this. It is simplest to set $\Phi_1 = 0$ since the issues really relate to the contemporaneous part of the system. Because we will be interested in a structural system it is necessary to differentiate between the VECM and the SVECM. To convert one to the other we pre-multiply (7.1) by a matrix of contemporaneous coefficients Φ_0 as follows:[1]

$$\begin{aligned} VECM &: \Delta\zeta_t = \alpha\beta'\zeta_{t-1} + \mathbf{e}_t \\ SVECM &: \mathbf{\Phi}_0\Delta\zeta_t = \mathbf{\Phi}_0\alpha\beta'\zeta_{t-1} + \Phi_0\mathbf{e}_t \\ &= \alpha^*\beta'\zeta_{t-1} + \varepsilon_t \\ &= \alpha^*\xi_{t-1} + \varepsilon_t \end{aligned}$$

7.3 SVAR Forms of the SVECM

7.3.1 Permanent and Transitory Shocks Only

We start with the SVECM where there are r transitory shocks and $n-r$ permanent shocks

$$SVECM : \mathbf{\Phi}_0\Delta\zeta_t = \alpha^*\xi_{t-1} + \varepsilon_t. \tag{7.2}$$

Consequently, there must be $n-r$ structural equations with permanent shocks, and we will choose these to be the first $n-r$ equations. They have the format

$$\Phi_{11}^0\Delta\zeta_{1t} + \Phi_{12}^0\Delta\zeta_{2t} = \alpha_1^*\xi_{t-1} + \varepsilon_{1t},$$

where ζ_{1t} is $(n-r)\times 1$ and ζ_{2t} is $r \times 1$. We now want to eliminate $\Delta\zeta_{2t}$ from the equations with the permanent shocks. This is done by using the co-integrating relations $\xi_t = \beta_1'\zeta_{1t} + \beta_2'\zeta_{2t}$. Inverting this equation yields

$$\zeta_{2t} = (\beta_2')^{-1}(\xi_t - \beta_1'\zeta_{1t}),$$

provided of course that β_2' is non-singular (an assumption that is commented on later in this sub-section).

Now this expression for ζ_{2t} can be used to eliminate it from the first block of $n-r$ equations in (7.2) yielding

$$\Phi_{11}^0\Delta\zeta_{1t} + \Phi_{12}^0(\beta_2')^{-1}(\Delta\xi_t - \beta_1'\Delta\zeta_{1t}) = \alpha_1^*\xi_{t-1} + \varepsilon_{1t}.$$

Thereafter, defining $A_{11}^0 = \Phi_{11}^0 - \Phi_{12}^0(\beta_2')^{-1}\beta_1'$ and $A_{12}^0 = \Phi_{12}^0(\beta_2')^{-1}$ this becomes

[1] Φ_0 is the equivalent of the A_0 in Chapter 4, but we wish to use A_0 for the later SVAR representation.

$$A_{11}^0 \Delta\zeta_{1t} + A_{12}^0 \Delta\xi_t = \alpha_1^* \xi_{t-1} + \varepsilon_{1t}. \tag{7.3}$$

There is a second block of r equations in the SVECM. Substituting for ζ_{2t} in these leaves

$$A_{21}^0 \Delta\zeta_{1t} + A_{22}^0 \Delta\xi_t = \alpha_2^* \xi_{t-1} + \varepsilon_{2t}, \tag{7.4}$$

which can be re-expressed as

$$\begin{aligned} A_{21}^0 \Delta\zeta_{1t} + A_{22}^0 \xi_t &= (A_{22}^0 + \alpha_2^*)\xi_{t-1} + \varepsilon_{2t} \\ &= A_{22}^1 \xi_{t-1} + \varepsilon_{2t}. \end{aligned} \tag{7.5}$$

Hence the SVAR will involve the $n-r$ variables $\Delta\zeta_{1t}$ and the r error correction terms ξ_t.

There are two points to note from the analysis above:

- The coefficients in the SVAR (A_j) are different to the SVECM (Φ_j).
- The shocks in the SVAR are the same as in the SVECM.

Consequently the issue is how to estimate the SVAR equations in (7.3) and (7.5). Pagan and Pesaran (2008) claimed that the knowledge that a particular structural equation (or equations) has a permanent shock (shocks) implies that the value of α^* in those structural equations will be zero. This means that these equations must have no lagged error correction terms present in them, i.e. ξ_{t-1} is missing from (7.3) because $\alpha_1^* = 0$. This feature frees up the lagged error correction terms to be used as instruments when estimating the parameters of the equations with permanent shocks. No such exclusion applies to the equations (7.5) and here one has to find instruments from another source. In the event that enough instruments are available to estimate (7.3) then $\hat{\varepsilon}_{1t}$ would qualify.

If one wants impulse response functions it is attractive to follow the approach of converting the SVECM to a SVAR. That is a relatively simple task to do so, since the co-integrating vector can be estimated separately and that enables the construction of the error correction terms.[2] Of course there is a question about choosing a set of $n-r$ variables ζ_{1t} from ζ_t. Provided β_2' is non-singular any $n-r$ variables can be chosen but, if it is not, the variables need to be selected in such a way as to make it non-singular. As an example, suppose there are three $I(1)$ variables with one co-integrating vector $\beta' = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}$. Then, if we choose the $n-r=2$ variables as ζ_{1t}, ζ_{2t}, we find that $\beta_2' = 0$. So it would be necessary to choose either $\{\zeta_{1t}, \zeta_{3t}\}$ or $\{\zeta_{2t}, \zeta_{3t}\}$ as the two variables. In these cases β_2' is either -1 or $+1$ and so non-singular.[3]

[2] The co-integrating vector is estimated super-consistently.
[3] Our thanks to Farshid Vahid for the example.

7.3.2 Permanent, Transitory and Mixed Shocks

Suppose now that the system contains $I(0)$ variables w_t as well as I(1) variables ζ_t. Then the most general SVECM would look like

$$\begin{aligned}\mathbf{\Phi}_0\Delta\zeta_t + \Psi w_t &= \alpha^*\xi_{t-1} + \varepsilon_t \\ Gw_t + H\Delta\zeta_t &= \delta\xi_{t-1} + \varepsilon_{3t},\end{aligned}$$

making the first $n-r$ equations

$$\Phi^0_{11}\Delta\zeta_{1t} + \Phi^0_{12}\Delta\zeta_{2t} + \Psi_{13}w_t = \alpha^*_1\xi_{t-1} + \varepsilon_{1t}.$$

Replacing $\Delta\zeta_{2t}$ using the co-integrating relation does not introduce any dependence upon w_t so the equivalent of (7.3) will be

$$A^0_{11}\Delta\zeta_{1t} + A^0_{12}\Delta\xi_t + \Psi_{13}w_t = \alpha^*_1\xi_{t-1} + \varepsilon_{1t}. \tag{7.6}$$

In the same way w_t is just added into (7.5) meaning that the SVAR will consist of $\Delta\zeta_{1t}, \xi_t$ and w_t.

The only change arises if we want the shocks coming from the introduction of the $I(0)$ variables to have transitory effects. It is clear that this will not happen with (7.6). Following the arguments in the preceding chapter, in order that the shock ε_{3t} has transitory effects it is necessary to specify (7.6) as

$$A^0_{11}\Delta\zeta_{1t} + A^0_{12}\Delta\xi_t + \Psi_{13}\Delta w_t = \alpha^*_1\xi_{t-1} + \varepsilon_{1t}. \tag{7.7}$$

Thus to estimate these equations we would use w_{t-1} as the instrument for Δw_t. All the other equations in the system feature the level of w_t. Hence when the extra shocks coming from the structural equations for w_t are made to have transitory effects it is necessary to specify the system so that Δw_t enters into those structural equations that have permanent shocks. If you want the ε_{3t} shocks to have permanent effects on the $I(1)$ variables then you can leave w_t in its level form those equations. Of course you will need to find an instrument for it when using the IV approach.

7.4 Example: Gali's (1999) Technology Shocks and Fluctuations Model

7.4.1 Nature of System and Restrictions Used

Gali (1999) has a five variable model consisting of labor productivity (x_t), the log of per capita hours or employment (n_t), the inflation rate ($\pi_t = \Delta\log p_t$), the nominal interest rate (i_t), and the growth rate of the money supply (Δm_t). All variables are taken to be $I(1)$ and there are two co-integrating relations, $\xi_{1t} = i_t - \pi_t$ and $\xi_{2t} = \Delta m_t - \pi_t$. Hence it is assumed that there will be three permanent shocks and two transitory shocks. By definition $\beta' = \begin{pmatrix} 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix}$ and the SVAR Gali uses contains $\Delta x_t, \Delta n_t, \Delta\pi_t, \xi_{1t}$ and ξ_{2t}. To convert to this form of SVAR from the SVECM we need $\beta_2' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ to be non-singular, which it is. For illustration it is assumed that the SVAR is of order 1, although it is of higher order in his paper and in our empirical implementation.

7.4.2 Estimation of the System

It is assumed that the equation with permanent technology shocks is the first. Given that fact it has a form like (7.3), namely

$$\begin{aligned} \Delta x_t &= \alpha_{12}^0 \Delta n_t + \alpha_{13}^0 \Delta\pi_t + \alpha_{14}^0 \Delta\xi_{1t} + \alpha_{15}^0 \Delta\xi_{2t} + \\ &\quad \alpha_{12}^1 \Delta n_{t-1} + \alpha_{13}^1 \Delta\pi_{t-1} + \alpha_{11}^1 \Delta x_{t-1} + \varepsilon_{1t}. \end{aligned} \tag{7.8}$$

To estimate (7.8) instruments for $\Delta n_t, \Delta\pi_t, \Delta\xi_{1t}$ and $\Delta\xi_{2t}$ are needed. Because ξ_{jt-1} $(j = 1, 2)$ are excluded from the equation these will provide two of the requisite instruments, but two more are needed. To get these Gali assumes that there are zero long-run effects of the non-technology permanent shocks ε_{2t} and ε_{3t} upon labor productivity. As seen in the previous chapter this means that $\alpha_{12}^0 = -\alpha_{12}^1$ and $\alpha_{13}^0 = -\alpha_{13}^1$. Together these restrictions mean that (7.8) becomes

$$\begin{aligned} \Delta x_t = {} & \alpha_{12}^0 \Delta^2 n_t + \alpha_{13}^0 \Delta^2 \pi_t + \alpha_{14}^0 \Delta\xi_{1t} + \\ & \alpha_{15}^0 \Delta\xi_{2t} + \alpha_{11}^1 \Delta x_{t-1} + \varepsilon_{1t}. \end{aligned} \tag{7.9}$$

Instruments to estimate this equation will be $\Delta n_{t-1}, \Delta\pi_{t-1}$, ξ_{1t-1} and ξ_{2t-1}. The IV parameter estimates are

$$\hat{\alpha}_{12}^0 = 1.1219, \hat{\alpha}_{13}^0 = -0.1834, \hat{\alpha}_{14}^0 = -1.1423, \hat{\alpha}_{15}^0 = 0.2352.$$

After this equation is estimated the residuals are a measure of the technology shock. As this is all Gali is interested in he does not estimate the remainder of the system, i.e. he is only estimating a *sub-set* of shocks and so does not have to specify the remainder of the model. So how then do we find the effects of technology on n_t etc.? Here he uses the assumption that technology shocks are uncorrelated with the others and we now need to see how that helps. Because the issue is a generic one we will first look at it in a general way, followed by Gali's case.

7.4.3 Recovering Impulse Responses to a Single Shock

The simplest way to see how to recover a sub-set of impulse responses without specifying the whole system is to note the relation between the VAR (or VECM) and structural errors viz.

$$e_t = A_0^{-1}\varepsilon_t = \bar{A}\varepsilon_t$$
$$so\ e_t = \sum_{j=1}^{n} \bar{A}_j \varepsilon_{jt},$$

where $\bar{A}_j$ is the j'th column of $\bar{A} = A_0^{-1}$. Looking at the $k'th$ VAR equation this will be

$$\begin{aligned} e_{kt} &= \sum_{j=1}^{n} \bar{a}_{kj}\varepsilon_{jt} && (7.10) \\ &= \bar{a}_{k1}\varepsilon_{1t} + \bar{a}_{k2}\varepsilon_{2t} + ... + \bar{a}_{kn}\varepsilon_{nt} \\ &= \bar{a}_{k1}\varepsilon_{1t} + \nu_{kt}. && (7.11) \end{aligned}$$

From this all that is needed to recover the impulse responses of all variables to the first shock would be the first column of $\bar{A}$, i.e. the elements $\bar{a}_{k1}$. These can be estimated by regressing the VAR (VECM) shocks $\hat{e}_{kt}$ on $\hat{\varepsilon}_{1t}$, because ε_{1t} is uncorrelated with all the other shocks ε_{jt} $(j \neq 1)$, and hence ν_{kt}. This is where the uncorrelatedness of shocks is important. With it all impulse responses to any structural shock can be recovered provided an estimate can be made of the requisite shock, i.e. ε_{1t}, since the VAR residuals $\hat{e}_t$ are available without specifying any structure.

In fact we don't need to run a regression to get the $\bar{a}_{k1}$. All that is needed is to add (7.9) on to the VAR equations for the remaining variables in the system so as to produce a combined SVAR/VAR structure. Now in the VAR module there is no allowance for such a hybrid structure so we need to use the SVAR commands modified in such a way as to allow a correlation between the errors of the VAR equations as well as a correlation with the structural equation error. To see how this can be done, the SVAR/VAR hybrid system for Gali's example looks like:

$$\begin{aligned} \Delta x_t &= \alpha_{12}^0 \Delta^2 n_t + \alpha_{13}^0 \Delta^2 \pi_t + \alpha_{14}^0 \Delta\xi_{1t} + && (7.12) \\ &\quad \alpha_{15}^0 \Delta\xi_{2t} + \alpha_{11}^1 \Delta x_{t-1} + \varepsilon_{1t} \\ \Delta n_t &= b_{21}^1 \Delta x_{t-1} + b_{22}^1 \Delta\pi_{t-1} + b_{23}^1 \xi_{1t-1} + b_{24}^1 \xi_{2t-1} + \bar{a}_{21}\varepsilon_{1t} + \nu_{2t} && (7.13) \\ \Delta\pi_t &= b_{31}^1 \Delta x_{t-1} + b_{32}^1 \Delta\pi_{t-1} + b_{33}^1 \xi_{1t-1} + b_{34}^1 \xi_{2t-1} + \bar{a}_{31}\varepsilon_{1t} + \nu_{3t} && (7.14) \\ \xi_{1t} &= b_{41}^1 \Delta x_{t-1} + b_{42}^1 \Delta\pi_{t-1} + b_{43}^1 \xi_{1t-1} + b_{44}^1 \xi_{2t-1} + \bar{a}_{41}\varepsilon_{1t} + \nu_{4t} && (7.15) \\ \xi_{2t} &= b_{51}^1 \Delta x_{t-1} + b_{52}^1 \Delta\pi_{t-1} + b_{53}^1 \xi_{1t-1} + b_{54}^1 \xi_{t-1} + \bar{a}_{51}\varepsilon_{1t} + \nu_{5t}, && (7.16) \end{aligned}$$

where ν_{jt} are combinations of $\varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}$ and ε_{5t}. These are uncorrelated with ε_{1t} but will generally be correlated with each other as they are in VAR equations. Hence, we need to allow for that in some way when estimating this augmented system.

EViews works with the (A, B) form $Ae_t = Bu_t$, where u_t are taken to be $i.i.d(0, 1)$ and uncorrelated with one another. So define ε_{1t} and ν_{jt} with the following relations

$$\begin{aligned}
\varepsilon_{1t} &= \delta_1 u_{1t} \\
\nu_{2t} &= \delta_2 u_{2t} \\
\nu_{3t} &= \delta_3 u_{2t} + \delta_4 u_{3t} \\
\nu_{4t} &= \delta_5 u_{2t} + \delta_6 u_{3t} + \delta_7 u_{4t} \\
\nu_{5t} &= \delta_8 u_{2t} + \delta_9 u_{3t} + \delta_{10} u_{4t} + \delta_{11} u_{5t}.
\end{aligned}$$

In this structure the shocks ν_{jt} are correlated with each other due to the presence of common elements, but they are uncorrelated with ε_{1t}. Hence this device also captures the nature of the combined SVAR/VAR system.

The complete system (7.12)-(7.16) can now be written in the A/B form by setting

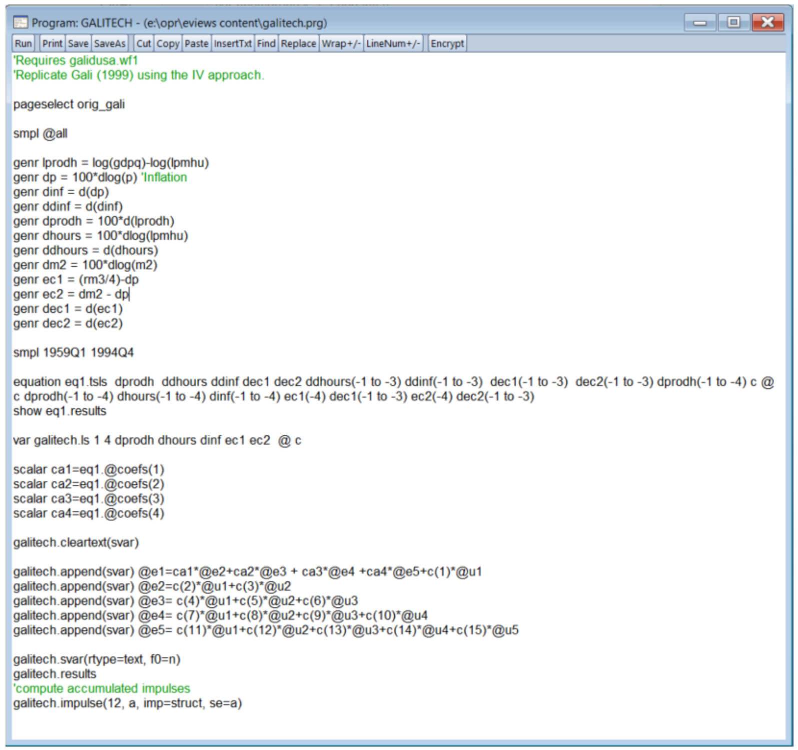
$$A = \begin{bmatrix} 1 & -\alpha_{12}^0 & -\alpha_{13}^0 & -\alpha_{14}^0 & -\alpha_{15}^0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} \delta_1 & 0 & 0 & 0 & 0 \\ \bar{a}_{21} & \delta_2 & 0 & 0 & 0 \\ \bar{a}_{31} & \delta_3 & \delta_4 & 0 & 0 \\ \bar{a}_{41} & \delta_5 & \delta_6 & \delta_7 & 0 \\ \bar{a}_{51} & \delta_8 & \delta_9 & \delta_{10} & \delta_{11} \end{bmatrix}$$

Because α_{ij}^0 have been estimated by IV it is only necessary to estimate B in EViews. There are fifteen unknown elements in B and fifteen parameters are in the covariance matrix of the reduced form VAR, so it is exactly identified.[4] *Galitech.prg* (see Figure 7.1) contains the code to estimate the model using the SVAR routine. Note that $\Delta x_t = dprodh, \Delta n_t = dhours, \Delta \pi_t = dinf, \xi_{1t} = ec1, \xi_{2t} = ec2$. Cumulative impulse responses are given in Figure 7.2. They match those of Gali (1999).

An alternative approach is to recognize that equations (7.12) - (7.13) can be estimated as a system using the FIML estimator and an unrestricted covariance matrix. To see this express the system as

$$\begin{aligned}
\Delta x_t &= \alpha_{12}^0 \Delta^2 n_t + \alpha_{13}^0 \Delta^2 \pi_t + \alpha_{14}^0 \Delta \xi_{1t} + \\
&\quad \alpha_{15}^0 \Delta \xi_{2t} + \alpha_{11}^1 \Delta x_{t-1} + \varepsilon_{1t} && (7.17) \\
\Delta n_t &= b_{21}^1 \Delta x_{t-1} + b_{22}^1 \Delta \pi_{t-1} + b_{23}^1 \xi_{1t-1} + b_{24}^1 \xi_{2t-1} + \vartheta_{2t} && (7.18) \\
\Delta \pi_t &= b_{31}^1 \Delta x_{t-1} + b_{32}^1 \Delta \pi_{t-1} + b_{33}^1 \xi_{1t-1} + b_{34}^1 \xi_{2t-1} + \vartheta_{3t} && (7.19) \\
\xi_{1t} &= b_{41}^1 \Delta x_{t-1} + b_{42}^1 \Delta \pi_{t-1} + b_{43}^1 \xi_{1t-1} + b_{44}^1 \xi_{2t-1} + \vartheta_{4t} && (7.20) \\
\xi_{2t} &= b_{51}^1 \Delta x_{t-1} + b_{52}^1 \Delta \pi_{t-1} + b_{53}^1 \xi_{1t-1} + b_{54}^1 \xi_{t-1} + \vartheta_{5t}, && (7.21)
\end{aligned}$$

[4]The four long-run restrictions mean that there are nineteen restrictions and so nineteen parameters can be estimated.




```
'Requires galidusa.wf1
'Replicate Gali (1999) using the IV approach.

pageselect orig_gali

smpl @all

genr lprodh = log(gdpq)-log(lpmhu)
genr dp = 100*dlog(p) 'Inflation
genr dinf = d(dp)
genr ddinf = d(dinf)
genr dprodh = 100*d(lprodh)
genr dhours = 100*dlog(lpmhu)
genr ddhours = d(dhours)
genr dm2 = 100*dlog(m2)
genr ec1 = (rm3/4)-dp
genr ec2 = dm2 - dp
genr dec1 = d(ec1)
genr dec2 = d(ec2)

smpl 1959Q1 1994Q4

equation eq1.tsls  dprodh  ddhours ddinf dec1 dec2 ddhours(-1 to -3) ddinf(-1 to -3)  dec1(-1 to -3)  dec2(-1 to -3) dprodh(-1 to -4) c @
c dprodh(-1 to -4) dhours(-1 to -4) dinf(-1 to -4) ec1(-4) dec1(-1 to -3) ec2(-4) dec2(-1 to -3)
show eq1.results

var galitech.ls 1 4 dprodh dhours dinf ec1 ec2  @ c

scalar ca1=eq1.@coefs(1)
scalar ca2=eq1.@coefs(2)
scalar ca3=eq1.@coefs(3)
scalar ca4=eq1.@coefs(4)

galitech.cleartext(svar)

galitech.append(svar) @e1=ca1*@e2+ca2*@e3 + ca3*@e4 +ca4*@e5+c(1)*@u1
galitech.append(svar) @e2=c(2)*@u1+c(3)*@u2
galitech.append(svar) @e3= c(4)*@u1+c(5)*@u2+c(6)*@u3
galitech.append(svar) @e4= c(7)*@u1+c(8)*@u2+c(9)*@u3+c(10)*@u4
galitech.append(svar) @e5= c(11)*@u1+c(12)*@u2+c(13)*@u3+c(14)*@u4+c(15)*@u5

galitech.svar(rtype=text, f0=n)
galitech.results
'compute accumulated impulses
galitech.impulse(12, a, imp=struct, se=a)
```

Figure 7.1: EViews Program *galitech.prg* to Estimate Gali's (1999) Technology Shocks SVAR

in which $cov(\varepsilon_{1t}, \vartheta_{2t}, \vartheta_{3t}, \vartheta_{4t}, \vartheta_{5t}) = BB'$ and is clearly non-diagonal in general. Ignoring the coefficients, the system has 19 unknown parameters[5] and hence is exactly identified.
The EViews SYSTEM object code required to estimate (7.17) to (7.21) using FIML is shown in Figure 7.3. As before, the long-run assumptions concerning the impact of money supply, money demand and aggregate demand on output growth are enforced by ensuring that the contemporaneous coefficients for the corresponding variables ($dhours$, $dinf$, $ec1$ and $ec2$) are equal but opposite in sign to the sum of the corresponding lag coefficients in the system. Estimating the model using the FIML estimator in EViews yields the results for Equation 7.17 shown in Figure 7.7. The implied FIML estimates for the contemporaneous parameters are identical to the IV/SVAR estimates, and hence re-produce the impulse response functions in Figure 7.2.

7.5 Example: Gali's 1992 IS/LM Model

7.5.1 Nature of System and Restrictions Used

Gali (1992) has a model with four $I(1)$ variables - the log of GNP at 1982 prices (y_t), the yield on three-month Treasury Bills (i_t), the growth in M1 (Δm_t)., and the inflation rate in the CPI (Δp_t). Hence $\zeta_t' = \begin{bmatrix} y_t & i_t & \Delta m_t & \Delta p_t \end{bmatrix}$. He indicates that there are two co-integrating vectors among these four variables. Therefore in this case $n = 4, r = 2$ and there are $n - r = 2$ permanent shocks and two transitory shocks. We take the permanent shocks as being those in the structural equations involving y_t and i_t. Gali works with an SVAR in terms of the variables $\Delta y_t, \Delta i_t, \xi_{1t} = i_t - \Delta p_t, \xi_{2t} = \Delta m_t - \Delta p_t$.

There are four shocks in the system which he calls (aggregate) supply (ε_{1t}), money supply (ε_{2t}), money demand (ε_{3t}) and an aggregate demand (IS) shock (ε_{4t}). Supply shocks can be taken to be permanent but one of the others must also be permanent with the remaining two being transitory. Given the fact that Gali works with Δi_t and treats i_t as $I(1)$ we will take the second permanent shock (ε_{2t}) as relating to money supply. Then he needs some extra restrictions to estimate the system and these are

- Both the money and aggregate demand shocks ε_{3t} and ε_{4t} are transitory.
- The money supply shock has a zero long-run effect on output.

[5] These are the four structural parameters, $\alpha_{12}^0, \alpha_{13}^0, \alpha_{14}^0 \; and \; \alpha_{15}^0$, and the 15 variance and covariance terms in BB'.

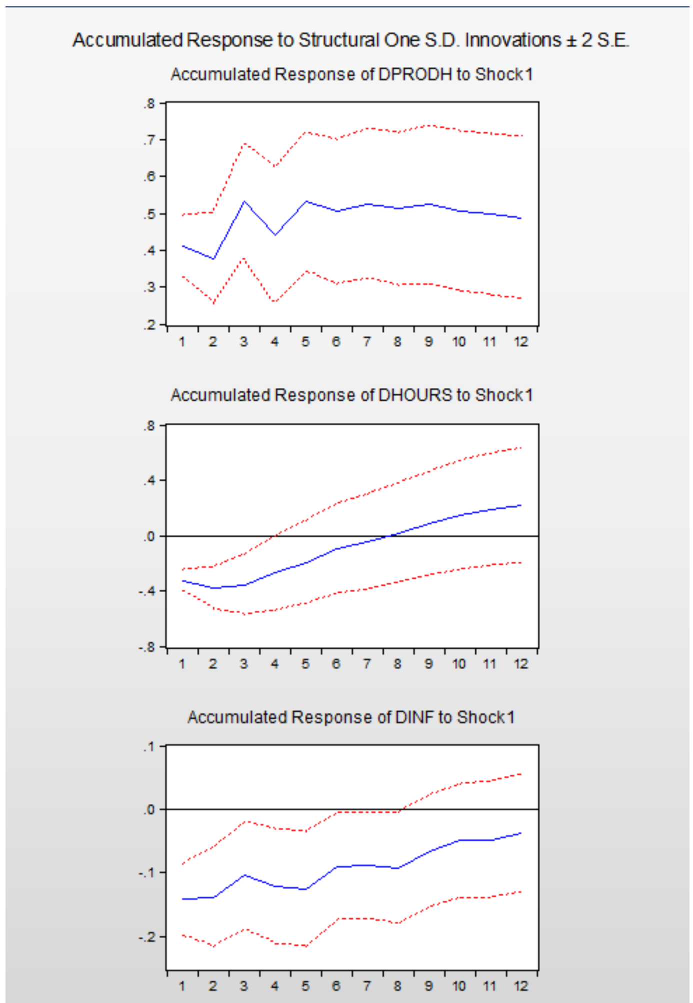


Figure 7.2: Accumulated Impulse Responses for Gali (1999) Featuring Technology Shocks

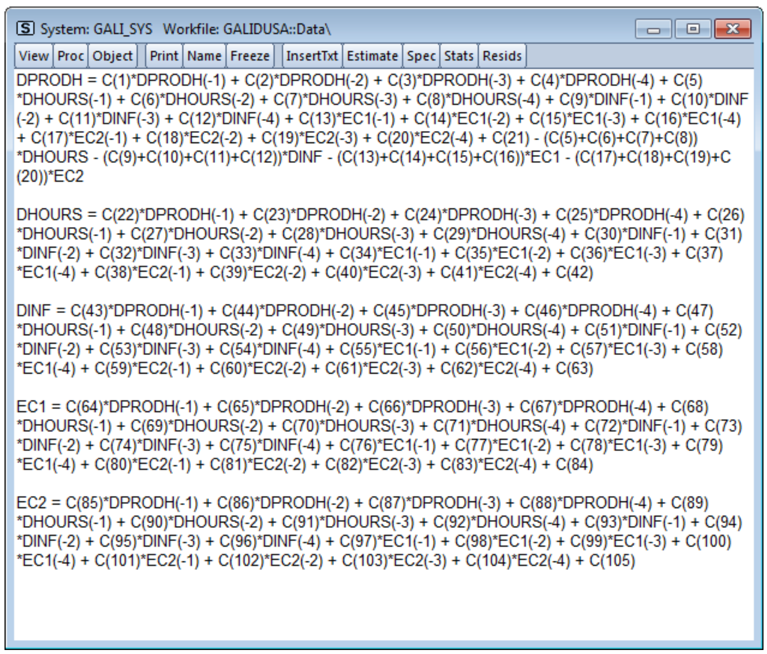


Figure 7.3: EViews SYSTEM Object Code (*gali_sys*) to Estimate (7.17) - (7.21) using FIML

System: GALI_SYS Workfile: GALIDUSA::orig_gali\

View | Proc | Object | Print | Name | Freeze | InsertTxt | Estimate | Spec | Stats | Resids

System: GALI_SYS
Estimation Method: Full Information Maximum Likelihood (OPG - BHHH / Marquardt steps)
Date: 02/27/16 Time: 06:27
Sample: 1960Q2 1994Q4
Included observations: 139
Total system (balanced) observations 695
Convergence achieved after 1847 iterations
Coefficient covariance computed using observed Hessian

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.474074	0.189301	-2.504339	0.0123
C(2)	0.022653	0.166329	0.136197	0.8917
C(3)	-0.367808	0.155044	-2.372273	0.0177
C(4)	-0.151900	0.143283	-1.060141	0.2891
C(5)	-0.666521	0.379005	-1.758609	0.0786
C(6)	-0.034594	0.222847	-0.155237	0.8766
C(7)	-0.103219	0.198514	-0.519956	0.6031
C(8)	-0.317574	0.187999	-1.689234	0.0912
C(9)	0.150735	0.715427	0.210692	0.8331
C(10)	-0.141292	0.605298	-0.233426	0.8154
C(11)	0.018790	0.813040	0.023111	0.9816
C(12)	0.155215	0.252655	0.614335	0.5390
C(13)	0.787146	1.657928	0.474777	0.6349
C(14)	-0.155936	1.094384	-0.142487	0.8867
C(15)	0.481038	1.243629	0.386802	0.6989
C(16)	0.030057	0.570350	0.052699	0.9580
C(17)	-0.132172	0.329424	-0.401222	0.6883
C(18)	-0.100866	0.242288	-0.416306	0.6772
C(19)	-0.024807	0.177993	-0.139371	0.8892
C(20)	0.022681	0.194801	0.116429	0.9073

Figure 7.4: FIML Estimates for Equation (7.17) (Partial Output)

Cast in terms of the long-run response matrices the assumptions as stated above imply the following structure:

$$
\mathrm{C}(1) \;=\; \begin{array}{c} \\ y_t \\ i_t \\ \Delta m_t \\ \Delta p_t \end{array}
\begin{array}{c} \begin{array}{cccc} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{3t} & \varepsilon_{4t} \end{array} \\
\left[\begin{array}{cccc} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \end{array}
$$

7.5.2 Estimation of the System

We begin with the first of Gali's equations (where the normalization is on y_t)

$$\Delta y_t = a_{12}^0 \Delta i_t + a_{13}^0 \Delta \xi_{1t} + a_{14}^0 \Delta \xi_{2t} + lags + \varepsilon_{1t},$$

where "lags" means lags of Δy_t etc. Since the first equation has a permanent shock we know from the Pagan-Pesaran (PP) result that only differences of the error correction terms are in this equation, so that the lagged error correction terms ξ_{1t-1} and ξ_{2t-1} provide instruments for $\Delta\xi_{1t}$ and $\Delta\xi_{2t}$. But one more instrument is needed for Δi_t. This is where Gali's assumption that money supply shocks (the shock in the Δi_t equation) have a zero long-run effect on output comes in. It implies that the coefficients of Δi_t and Δi_{t-1} are equal and opposite so that the equation becomes

$$\Delta y_t = a_{12}^0 \Delta^2 i_t + a_{13}^0 \Delta \xi_{1t} + a_{14}^0 \Delta \xi_{2t} + lags + \varepsilon_{1t}, \tag{7.22}$$

yielding the third instrument Δi_{t-1}. This equation can then be estimated and residuals $\hat{\varepsilon}_{1t}$ recovered.

Now because the second equation also has a permanent shock, it follows that it should have the form

$$\Delta i_t = a_{21}^0 \Delta y_t + a_{23}^0 \Delta \xi_{1t} + a_{24}^0 \Delta \xi_{2t} + lags + \varepsilon_{2t}. \tag{7.23}$$

Again the lagged error correction terms can be used as instruments. This leaves us the task of finding one instrument for Δy_t. But that is available from the residuals $\hat{\varepsilon}_{1t}$. Hence (7.23) can be estimated using the instruments provided by the assumption of cointegration.

Gali does not do this. The equation he actually estimates is of the form (see Pagan and Robertson (1998 p. 213))

$$\Delta i_t = \gamma_{21}^0 \Delta y_t + \gamma_{23}^0 \xi_{1t} + \gamma_{24}^0 \xi_{2t} + lags + \varepsilon_{2t}. \tag{7.24}$$

This differs from the correct structure (7.23) because it involves the *level* of the co-integrating errors and *not the changes*. Because of this difference Gali does not treat the lagged error correction terms available as instruments and is therefore forced to impose two short-run restrictions. He lists three possible short-run restrictions - what he calls $R4$, $R5$ and $R6$. $R4$ and $R5$ imply no contemporaneous effects of money supply ($R4$) and demand shocks ($R5$) on output implying that the (1,2) and (1,3) elements of A_0^{-1} are zero. $R6$ says that contemporaneous prices don't enter the money supply rule, meaning $\gamma_{23}^0+\gamma_{24}^0 = 0$.

Now the first of these ($R4$) means that the VAR output growth equation errors e_{1t} does not involve ε_{2t} and so the VAR residuals $\hat{e}_{1t}$ can be used as an instrument in the second equation for i_t. $R5$ does not deliver any usable instruments for this equation (since it says e_{2t} is uncorrelated with ε_{3t}) and therefore $R6$ is needed, i.e. $\gamma_{23}^0+\gamma_{24}^0 = 0$. As emphasized above neither of these is actually required because ξ_{1t-1} and ξ_{2t-1} provide the instruments, given *his* co-integrating assumptions. Recall that there must be two permanent shocks in the system (since the number is $n-r$ and he has set $r=2$). If one follows through on the implications of Gali's $I(1)$ and co-integrating assumptions it is not necessary to use short-run restrictions to estimate this equation. Gali treats the shock of the second equation as transitory when in fact it is not transitory. This may stem from a confusion between the stochastic nature of the shock and its effects. The shock ε_{2t} is an $I(0)$ process but it has a permanent effect upon i_t.

Anyway, if one follows Gali's approach of using R6, the equation estimated would be

$$\Delta i_t = \gamma_{21}^0 \Delta y_t + \gamma_{23}^0(\xi_{1t}-\xi_{2t}) + lags + \varepsilon_{2t}, \tag{7.25}$$

where $\hat{\varepsilon}_{1t}$ and $\hat{e}_{1t}$ are used as instruments.

Moving on to the third equation it will be

$$\xi_{1t} = a_{31}^0 \Delta y_t + a_{32}^0 \Delta i_t + a_{34}^0 \xi_{2t} + lags + \varepsilon_{3t}$$

.Now the residuals $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ are available as instruments for Δy_t and Δi_t but another is needed for ξ_{2t}. Here a short-run restriction is needed. If one follows what Gali did then the logical restriction is that money demand shocks have no contemporaneous effect on output. This means that the VAR equation for Δy_t will have a shock that does not include ε_{3t}, a result established earlier in Chapter 4. Hence e_{1t} can be used as an instrument in this equation as well.

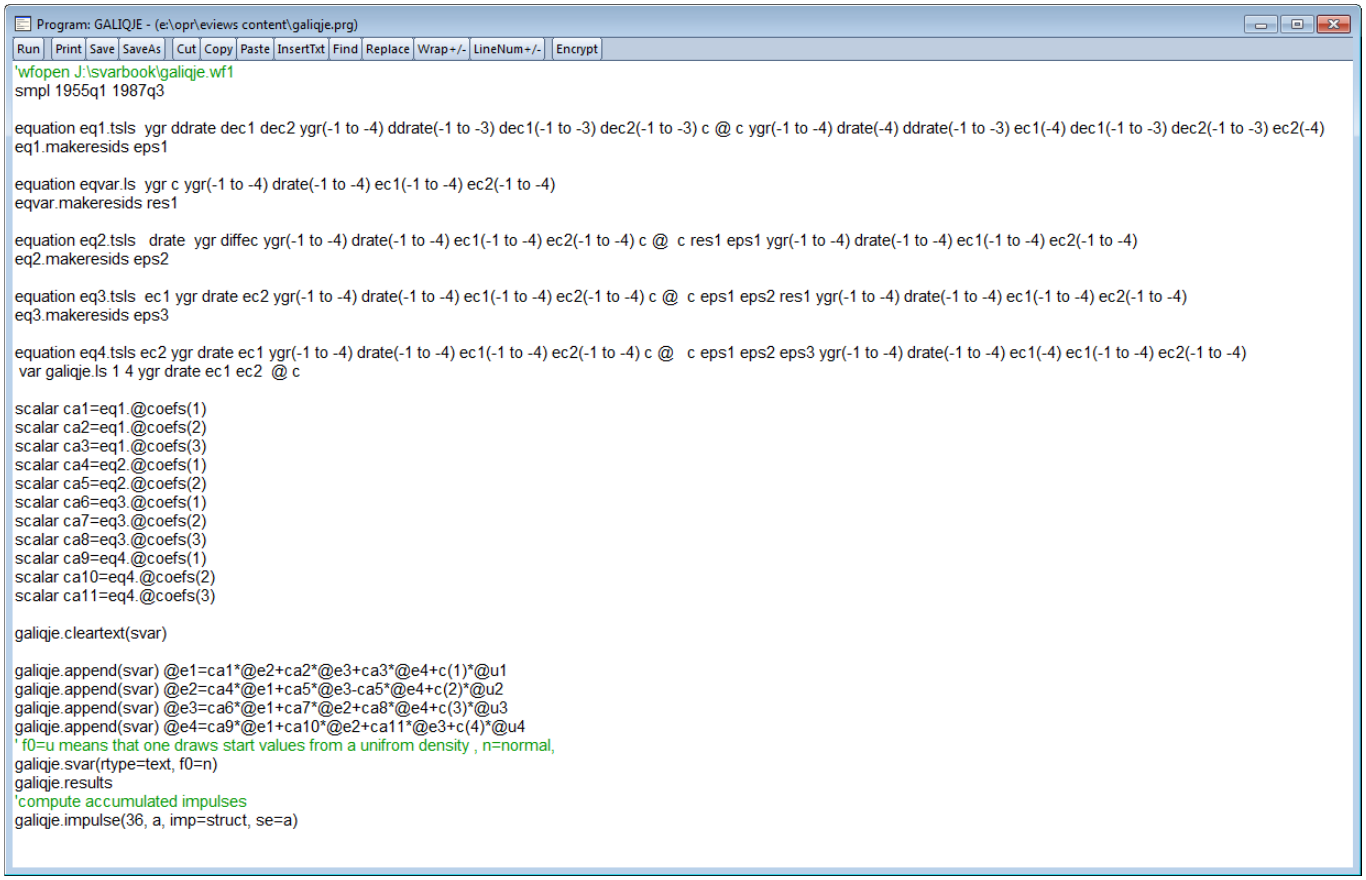
Once this equation is estimated, $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ and $\hat{\varepsilon}_{3t}$ are available as instruments to estimate the remaining equation in the system

$$\xi_{2t} = a_{41}^0 \Delta y_t + a_{42}^0 \Delta i_t + a_{43}^0 \xi_{1t} + lags + \varepsilon_{4t}. \tag{7.26}$$

Program *galiqje.prg* in Figure 7.5 contains the code to estimate the model in this way. In the program a VAR(4) is estimated in line with what Gali did. The correspondence between model variables and data is:

$$\Delta y_t = ygr, \Delta i_t = drate, \xi_{1t} = ec1, \xi_{2t} = ec2 \text{ and } \xi_{1t}-\xi_{2t} = diffec.$$

Figure 7.6 shows the resulting parameter estimates, and Figures 7.7 and 7.8 show the responses of the level of GNP and the interest rate to the four shocks in the system. These closely follow what is in Gali (1992). Looking at the impulse



Program: GALIQJE - (e:\opr\eviews content\galiqje.prg)

Run | Print | Save | SaveAs | Cut | Copy | Paste | InsertTxt | Find | Replace | Wrap+/- | LineNum+/- | Encrypt

```
'wfopen J:\svarbook\galiqje.wf1
smpl 1955q1 1987q3

equation eq1.tsls  ygr ddrate dec1 dec2 ygr(-1 to -4) ddrate(-1 to -3) dec1(-1 to -3) dec2(-1 to -3) c @ c ygr(-1 to -4) drate(-4) ddrate(-1 to -3) ec1(-4) dec1(-1 to -3) dec2(-1 to -3) ec2(-4)
eq1.makeresids eps1

equation eqvar.ls  ygr c ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
eqvar.makeresids res1

equation eq2.tsls   drate  ygr diffec ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4) c @  c res1 eps1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
eq2.makeresids eps2

equation eq3.tsls  ec1 ygr drate ec2 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4) c @  c eps1 eps2 res1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
eq3.makeresids eps3

equation eq4.tsls ec2 ygr drate ec1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4) c @   c eps1 eps2 eps3 ygr(-1 to -4) drate(-1 to -4) ec1(-4) ec1(-1 to -4) ec2(-1 to -4)
 var galiqje.ls 1 4 ygr drate ec1 ec2  @ c

scalar ca1=eq1.@coefs(1)
scalar ca2=eq1.@coefs(2)
scalar ca3=eq1.@coefs(3)
scalar ca4=eq2.@coefs(1)
scalar ca5=eq2.@coefs(2)
scalar ca6=eq3.@coefs(1)
scalar ca7=eq3.@coefs(2)
scalar ca8=eq3.@coefs(3)
scalar ca9=eq4.@coefs(1)
scalar ca10=eq4.@coefs(2)
scalar ca11=eq4.@coefs(3)

galiqje.cleartext(svar)

galiqje.append(svar) @e1=ca1*@e2+ca2*@e3+ca3*@e4+c(1)*@u1
galiqje.append(svar) @e2=ca4*@e1+ca5*@e3-ca5*@e4+c(2)*@u2
galiqje.append(svar) @e3=ca6*@e1+ca7*@e2+ca8*@e4+c(3)*@u3
galiqje.append(svar) @e4=ca9*@e1+ca10*@e2+ca11*@e3+c(4)*@u4
' f0=u means that one draws start values from a unifrom density , n=normal,
galiqje.svar(rtype=text, f0=n)
galiqje.results
'compute accumulated impulses
galiqje.impulse(36, a, imp=struct, se=a)
```

Figure 7.5: EViews Program *galiqje.prg* to Estimate Gali's (1992) IS-LM Model

Var: GALIQJE Workfile: GALIQJE::GALI_92\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 02/08/16 Time: 12:16
Sample: 1955Q1 1987Q3
Included observations: 131
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 8 iterations
Structural VAR is over-identified (6 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@E1=CA1*@E2+CA2*@E3+CA3*@E4+C(1)*@U1
@E2=CA4*@E1+CA5*@E3-CA5*@E4+C(2)*@U2
@E3=CA6*@E1+CA7*@E2+CA8*@E4+C(3)*@U3
@E4=CA9*@E1+CA10*@E2+CA11*@E3+C(4)*@U4
where
@e1 represents YGR residuals
@e2 represents DRATE residuals
@e3 represents EC1 residuals
@e4 represents EC2 residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-1.070711	0.066149	-16.18641	0.0000
C(2)	-1.699136	0.104973	-16.18641	0.0000
C(3)	6.515471	0.402527	16.18641	0.0000
C(4)	8.202743	0.506767	16.18641	0.0000

Log likelihood	-955.1849			
LR test for over-identification:				
Chi-square(6)	0.000000		Probability	1.0000

Estimated A matrix:

1.000000	-0.655408	0.293322	-0.086442
0.239160	1.000000	0.301674	-0.301674
-1.557400	4.985212	1.000000	1.213470
4.654613	7.582059	-3.393284	1.000000

Estimated B matrix:

1.070711	0.000000	0.000000	0.000000
0.000000	1.699136	0.000000	0.000000
0.000000	0.000000	6.515471	0.000000
0.000000	0.000000	0.000000	8.202743

Figure 7.6: IV/SVAR Estimates for Gali's (1992) IS-LM Model

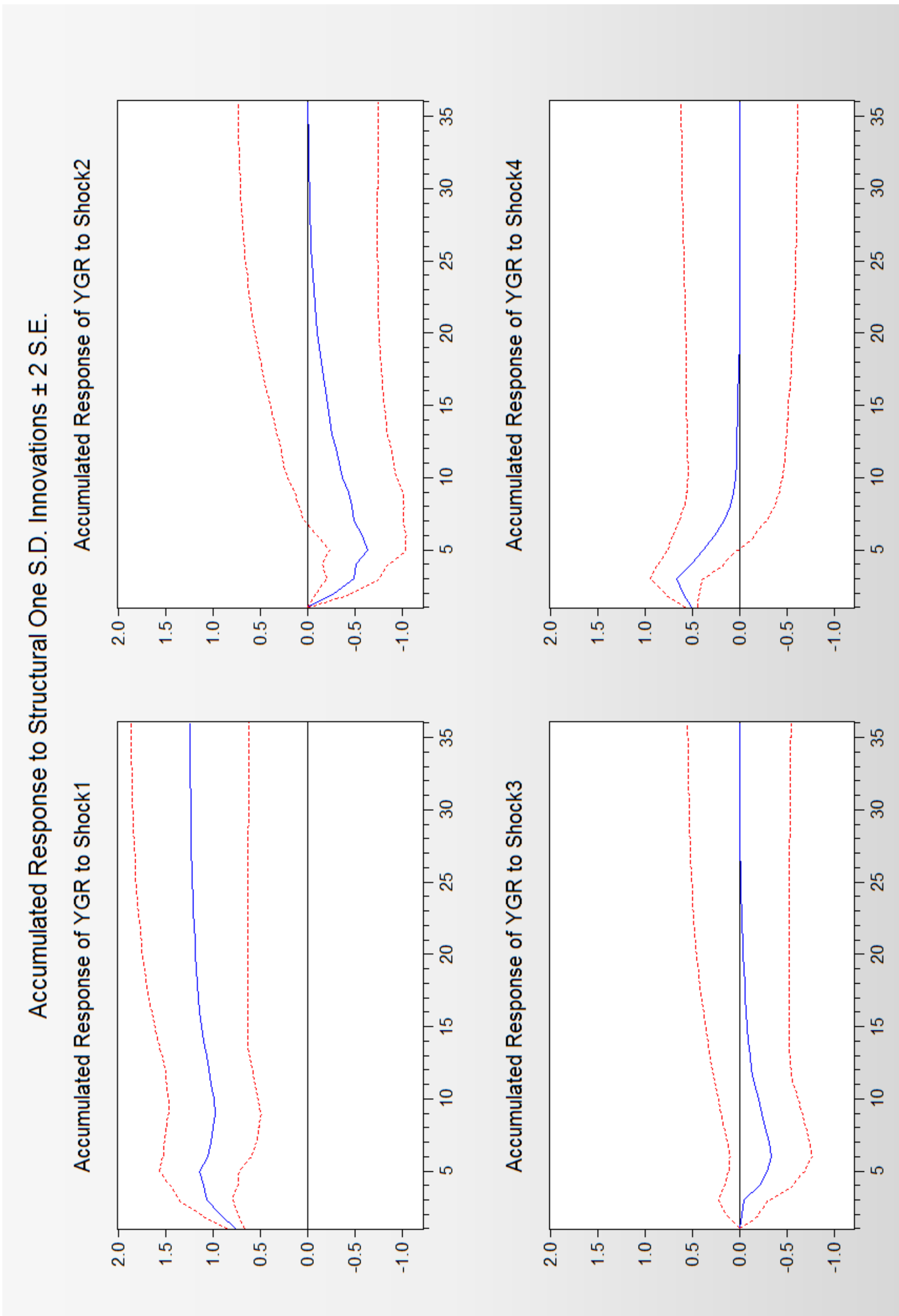


Figure 7.7: Impulse Responses of GNP using Galí's (1992) Restrictions

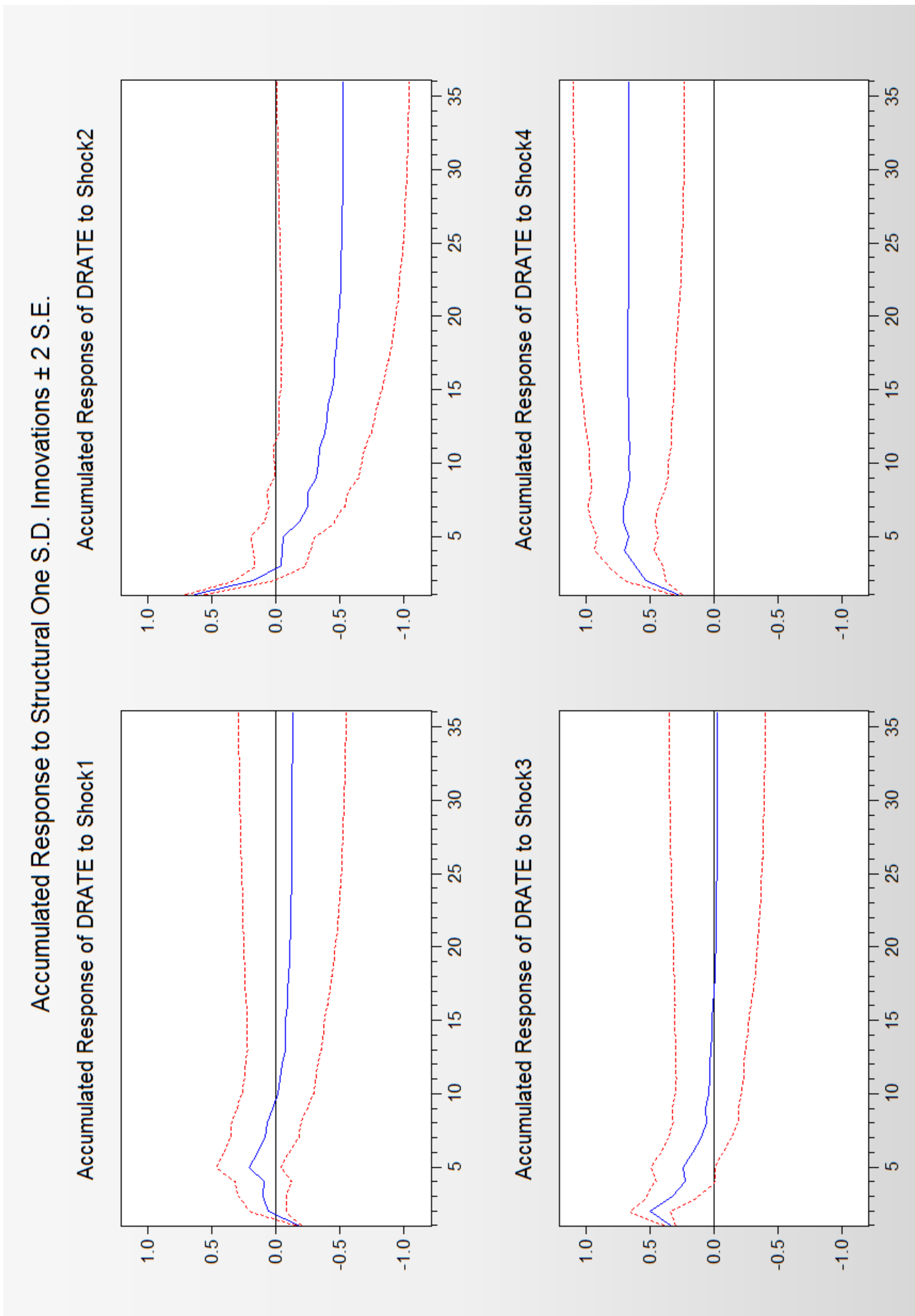


Figure 7.8: Impulse Responses for Δi_t using Galí's (1992) Restrictions

responses we see that a (negative) money supply shock has a negative effect on the level of output at all lags but goes to zero in the long run (as imposed). So there is no output puzzle. But we also notice that a demand shock has a long-run effect on the level of interest rates (Figure 7.8), which is unsatisfactory. It arises from the fact that the restriction that the third and fourth shocks of the system are transitory has not been imposed. One has to *impose* this effect upon the interest rate equation - it does not occur naturally.[6]

Estimating the model using FIML reveals implicit cross-equation constraints that are imposed automatically using the IV approach. Solving for the reduced form of Gali's system and collecting like terms, it can be shown that

$$\begin{aligned}\frac{\partial \Delta y_t}{\Delta i_t} &= \frac{1}{f}\{a_{12}^0 + a_{13}^0 a_{32} + a_{14}^0 a_{42}^0 + a_{13}^0 a_{34}^0 a_{42}^0 \\ &\quad + a_{14}^0 a_{32}^0 a_{43}^0 - a_{12}^0 a_{34}^0 a_{43}^0\} \\ \frac{\partial \Delta y_t}{\Delta \xi_{1t}} &= \frac{1}{f}\{(a_{13}^0 + a_{14}^0 a_{43}^0) + [\gamma_{23}^0(a_{12}^0 + a_{13}^0 a_{42}^0 + a_{14}^0 a_{42}^0 - a_{12}^0 a_{43}^0)]\},\end{aligned} \tag{7.27}$$

where f is the determinant of the contemporaneous coefficient matrix

$$A^0 = \begin{bmatrix} 1 & -a_{12}^0 & -a_{13}^0 & -a_{14}^0 \\ -\gamma_{21}^0 & 1 & -\gamma_{23}^0 & \gamma_{23}^0 \\ -a_{31}^0 & -a_{32}^0 & 1 & -a_{34}^0 \\ -a_{41}^0 & -a_{42}^0 & -a_{43}^0 & 1 \end{bmatrix}.$$

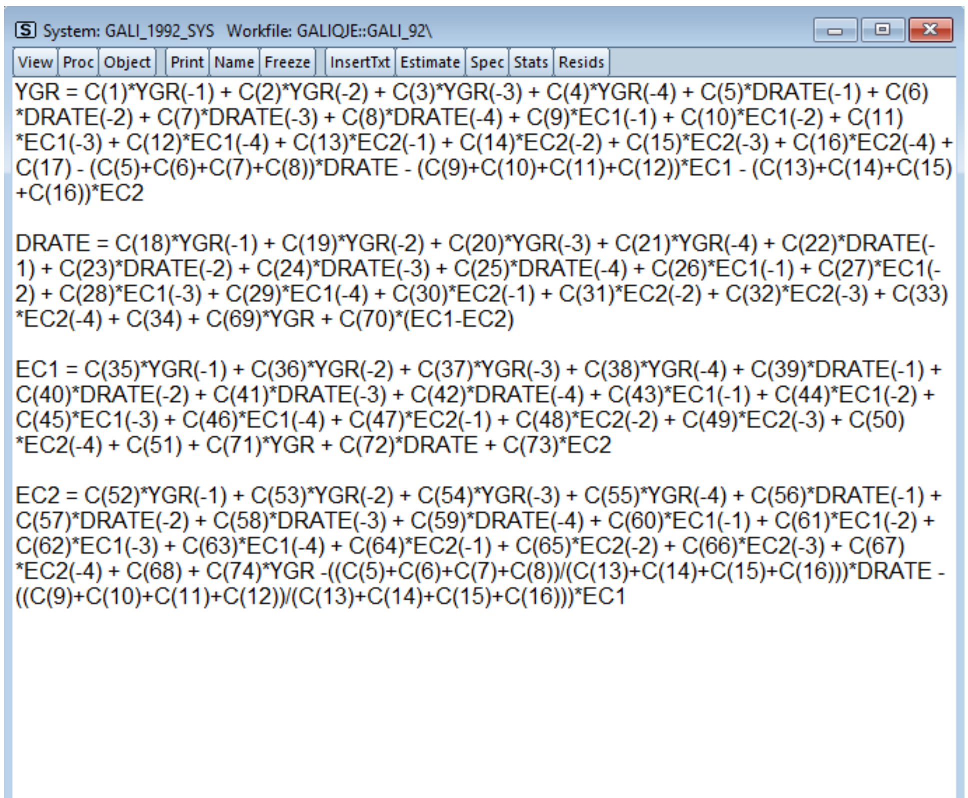
Note that to enforce Gali's short-run restrictions (i.e., $\frac{\partial \Delta y_t}{\Delta i_t} = 0$ and $\frac{\partial \Delta y_t}{\Delta \xi_{1t}} = 0$) it is sufficient to set $a_{42}^0 = -\left(\frac{a_{12}^0}{a_{14}^0}\right)$ and $a_{43}^0 = -\left(\frac{a_{13}^0}{a_{14}^0}\right)$. With these restrictions, the value of γ_{23}^0 does not affect $\frac{\partial \Delta y_t}{\Delta \xi_{1t}}$, since $(a_{12}^0 + a_{13}^0 a_{42}^0 + a_{14}^0 a_{42}^0 - a_{12}^0 a_{43}^0) = 0$. Moreover, substituting (7.26) into the output growth equation (7.22) gives

$$\begin{aligned}\Delta y_t &= (a_{12}^0 + a_{14}^0 a_{42}^0)\Delta^2 i_t + (a_{13}^0 + a_{14}^0 a_{43}^0)\Delta\xi_{1t} \\ &\quad + a_{14}^0 a_{41}^0 \Delta^2 y_t + lags + \varepsilon_{1t} + a_{14}^0 \Delta\varepsilon_{4t}.\end{aligned} \tag{7.28}$$

which becomes ,under the stated parameter restrictions. Output growth (Δy_t) is clearly independent of ε_{2t} and ε_{3t} with these restrictions imposed, and aggregate demand shocks (ε_{4t}) do not influence output in the long run since the long-run variance of $\Delta\varepsilon_{4t}$ is zero.

The EViews system object code to estimate Gali's model using FIML is given in Figure 7.9 and the final estimates for the contemporaneous parameters are shown in Figure 7.10. They were obtained assuming a diagonal covariance structure. Note that $a_{12} = -(C(5) + C(6) + C(7) + C(8)) = 0.65541$ with a standard error of 0.3356 using the delta method. This matches the estimate for a_{12}^0 obtained using the IV/SVAR approach (see Figure 7.6).

[6]There seem to be some weak instrument issues in Gali's estimation so that the EViews standard errors for impulse responses may not be very reliable. Pagan and Robertson (1998) found by simulation that multi-modal densities were likely for Gali's estimators rather than normality.




```
YGR = C(1)*YGR(-1) + C(2)*YGR(-2) + C(3)*YGR(-3) + C(4)*YGR(-4) + C(5)*DRATE(-1) + C(6)
*DRATE(-2) + C(7)*DRATE(-3) + C(8)*DRATE(-4) + C(9)*EC1(-1) + C(10)*EC1(-2) + C(11)
*EC1(-3) + C(12)*EC1(-4) + C(13)*EC2(-1) + C(14)*EC2(-2) + C(15)*EC2(-3) + C(16)*EC2(-4) +
C(17) - (C(5)+C(6)+C(7)+C(8))*DRATE - (C(9)+C(10)+C(11)+C(12))*EC1 - (C(13)+C(14)+C(15)
+C(16))*EC2

DRATE = C(18)*YGR(-1) + C(19)*YGR(-2) + C(20)*YGR(-3) + C(21)*YGR(-4) + C(22)*DRATE(-
1) + C(23)*DRATE(-2) + C(24)*DRATE(-3) + C(25)*DRATE(-4) + C(26)*EC1(-1) + C(27)*EC1(-
2) + C(28)*EC1(-3) + C(29)*EC1(-4) + C(30)*EC2(-1) + C(31)*EC2(-2) + C(32)*EC2(-3) + C(33)
*EC2(-4) + C(34) + C(69)*YGR + C(70)*(EC1-EC2)

EC1 = C(35)*YGR(-1) + C(36)*YGR(-2) + C(37)*YGR(-3) + C(38)*YGR(-4) + C(39)*DRATE(-1) +
C(40)*DRATE(-2) + C(41)*DRATE(-3) + C(42)*DRATE(-4) + C(43)*EC1(-1) + C(44)*EC1(-2) +
C(45)*EC1(-3) + C(46)*EC1(-4) + C(47)*EC2(-1) + C(48)*EC2(-2) + C(49)*EC2(-3) + C(50)
*EC2(-4) + C(51) + C(71)*YGR + C(72)*DRATE + C(73)*EC2

EC2 = C(52)*YGR(-1) + C(53)*YGR(-2) + C(54)*YGR(-3) + C(55)*YGR(-4) + C(56)*DRATE(-1) +
C(57)*DRATE(-2) + C(58)*DRATE(-3) + C(59)*DRATE(-4) + C(60)*EC1(-1) + C(61)*EC1(-2) +
C(62)*EC1(-3) + C(63)*EC1(-4) + C(64)*EC2(-1) + C(65)*EC2(-2) + C(66)*EC2(-3) + C(67)
*EC2(-4) + C(68) + C(74)*YGR -((C(5)+C(6)+C(7)+C(8))/(C(13)+C(14)+C(15)+C(16)))*DRATE -
((C(9)+C(10)+C(11)+C(12))/(C(13)+C(14)+C(15)+C(16)))*EC1
```

Figure 7.9: EViews SYSTEM Object Code (*gali_sys_1992*) to Estimate Gali's (1992) IS-LM Model

Parameter	Estimate	Standard Error	Sums/SE
C(5)	0.071914503	0.161876346	
C(6)	-0.048090118	0.150753672	
C(7)	-0.279488962	0.186331812	
C(8)	-0.399743613	0.136628237	0.655408
C(9)	0.044638472	0.073960125	(0.3356)
C(10)	0.093297416	0.08482032	
C(11)	0.159056097	0.076886262	
C(12)	-0.003669792	0.074436073	-0.29332
C(13)	-0.058223397	0.05750393	(-0.1937)
C(14)	0.004606155	0.035587852	
C(15)	-0.084972421	0.039955643	
C(16)	0.052147687	0.044469485	0.086442
C(69)	-0.239160488	0.671165099	(0.0603)
C(70)	-0.301673681	0.798521369	
C(71)	1.557400073	3.171887201	
C(72)	-4.985212364	9.592994904	
C(73)	-1.213469608	2.884438872	
C(74)	-4.6546134	3.89024338	

Figure 7.10: FIML Estimates (Diagonal Covariance) for Gali's (1992) IS-LM Model

Now consider what happens when the long-run restrictions implied by Gali's cointegration assumptions above are imposed. To recap, restrictions R4 and R6 are no longer necessary, since they are replaced by the adding up constraints on a_{23}^0 and a_{24}^0 in equation 7.23. The commands to implement the IV/SVAR approach are given in the program *galiqje_alt.prg* (Figure 7.11) and cumulative impulse responses are given in Figures 7.13 and 7.14.[7] Notice that an aggregate demand shock does not have a long-run effect on the nominal interest rate under the cointegration restrictions.

[7]The system object code for this case may be found in *gali_sys_alt* in the *galiqje.wk1* workfile. The MLE implementation using *optimize()* can be found in *gali_alt_mle.prg* in the MLE sub-directory. The system requires one additional restriction to achieve exact identification. We follow Gali and assume that R5 holds, namely $\frac{\partial \Delta y_t}{\Delta \xi_{1t}} = 0$. It can be shown that this requires $a_{43}^0 = (a_{13}^0(1.0 - a_{24}^0 a_{42}^0) - a_{23}^0(a_{12}^0 - a_{14}^0 a_{42}^0))/(a_{14}^0 - a_{12}^0 a_{24}^0)$. The resulting FIML estimates match the IV/SVAR estimates.

Program: GALIQJE_ALT - (e:\opr\eviews content\galiqje_alt.prg)

Run | Print | Save | SaveAs | Cut | Copy | Paste | InsertTxt | Find | Replace | Wrap+/- | LineNum+/- | Encrypt

```
'wfopen J:\svarbook\galiqje.wf1
pageselect gali_alt
smpl 1955Q1 1987Q3

equation alt_eq1.tsls ygr ddrate dec1 dec2 ygr(-1 to -4) ddrate(-1 to -3) dec1(-1 to -3) dec2(-1 to -3) c @ c ygr(-1 to -4) drate(-4) ddrate(-1 to -3) ec1(-4) dec1(-1 to -3) dec2(-1 to -3) ec2(-4)
alt_eq1.makeresids eps1

equation alt_eqvar.ls ygr c ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
alt_eqvar.makeresids res1

equation alt_eq2.tsls drate ygr dec1 dec2  ygr(-1 to -4) drate(-1 to -4) dec1(-1 to -3) dec2(-1 to -3) c @  c  eps1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
alt_eq2.makeresids eps2lr

equation alt_eq3.tsls ec1 ygr drate ec2 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4) c @  c eps1 eps2lr res1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4)
alt_eq3.makeresids eps3

equation alt_eq4.tsls ec2 ygr drate ec1 ygr(-1 to -4) drate(-1 to -4) ec1(-1 to -4) ec2(-1 to -4) c @   c eps1 eps2lr eps3 ygr(-1 to -4) drate(-1 to -4) ec1(-4) ec1(-1 to -4) ec2(-1 to -4)

var galiqjealt.ls 1 4 ygr drate ec1 ec2  @ c

scalar alt_ca1=alt_eq1.@coefs(1)
scalar alt_ca2=alt_eq1.@coefs(2)
scalar alt_ca3=alt_eq1.@coefs(3)

scalar alt_ca4=alt_eq2.@coefs(1)
scalar alt_ca5=alt_eq2.@coefs(2)
scalar alt_ca6=alt_eq2.@coefs(3)

scalar alt_ca7=alt_eq3.@coefs(1)
scalar alt_ca8=alt_eq3.@coefs(2)
scalar alt_ca9=alt_eq3.@coefs(3)

scalar alt_ca10=alt_eq4.@coefs(1)
scalar alt_ca11=alt_eq4.@coefs(2)
scalar alt_ca12=alt_eq4.@coefs(3)

galiqjealt.cleartext(svar)

galiqjealt.append(svar) @e1=alt_ca1*@e2+alt_ca2*@e3+alt_ca3*@e4+c(1)*@u1
galiqjealt.append(svar) @e2=alt_ca4*@e1+alt_ca5*@e3+alt_ca6*@e4+c(2)*@u2
galiqjealt.append(svar) @e3=alt_ca7*@e1+alt_ca8*@e2+alt_ca9*@e4+c(3)*@u3
galiqjealt.append(svar) @e4=alt_ca10*@e1+alt_ca11*@e2+alt_ca12*@e3+c(4)*@u4

' f0=u means that one draws start values from a unifrom density , n=normal,
galiqjealt.svar(rtype=text, f0=n)
galiqjealt.results
'compute accumulated impulses
galiqjealt.impulse(36, a, imp=struct, se=a)
```

Figure 7.11: EViews Program to Estimate Gali's IS-LM Model Using Alternative Restrictions

Var: GALIQJEALT Workfile: GALIQJE::GALI_ALT\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 02/28/16 Time: 07:56
Sample: 1955Q1 1987Q3
Included observations: 131
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 8 iterations
Structural VAR is over-identified (6 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@E1=ALT_CA1*@E2+ALT_CA2*@E3+ALT_CA3*@E4+C(1)*@U1
@E2=ALT_CA4*@E1+ALT_CA5*@E3+ALT_CA6*@E4+C(2)*@U2
@E3=ALT_CA7*@E1+ALT_CA8*@E2+ALT_CA9*@E4+C(3)*@U3
@E4=ALT_CA10*@E1+ALT_CA11*@E2+ALT_CA12*@E3+C(4)*@U4
where
@e1 represents YGR residuals
@e2 represents DRATE residuals
@e3 represents EC1 residuals
@e4 represents EC2 residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.070711	0.066149	16.18641	0.0000
C(2)	-3.007241	0.185788	-16.18641	0.0000
C(3)	6.945797	0.429113	16.18641	0.0000
C(4)	18.10785	1.118707	16.18641	0.0000

Log likelihood -955.1849
LR test for over-identification:
Chi-square(6) 0.000000 Probability 1.0000

Estimated A matrix:

1.000000	-0.655408	0.293322	-0.086442
1.095718	1.000000	-1.320536	0.600580
-1.709184	5.093092	1.000000	1.356424
-8.211299	-19.31645	2.036959	1.000000

Estimated B matrix:

1.070711	0.000000	0.000000	0.000000
0.000000	3.007241	0.000000	0.000000
0.000000	0.000000	6.945797	0.000000
0.000000	0.000000	0.000000	18.10785

Figure 7.12: IV/SVAR Estimates for Gali (1992) Using Alternative Restrictions

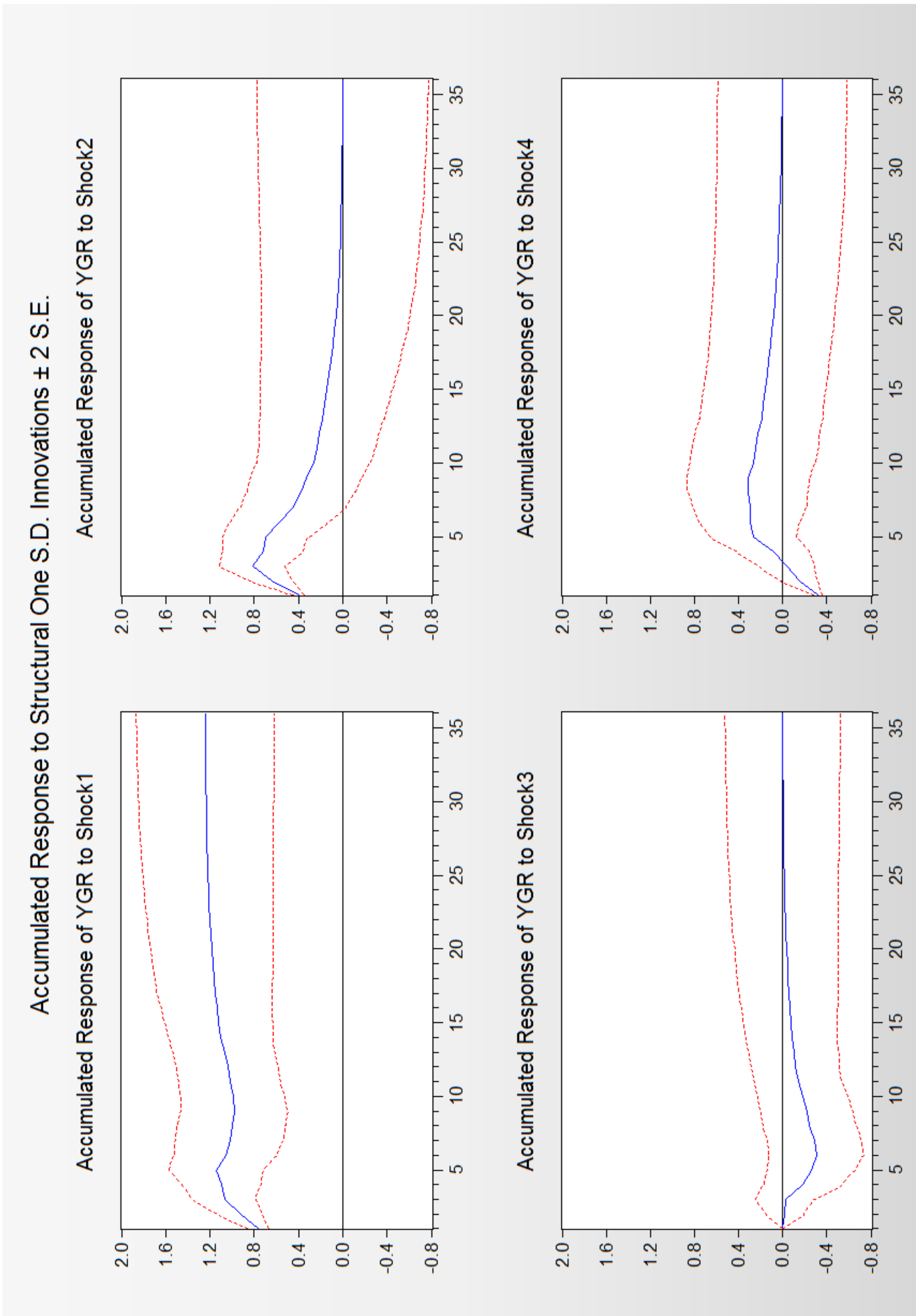


Figure 7.13: Accumulated Impulse Responses of GNP for Gali (1992) using Alternative Restrictions

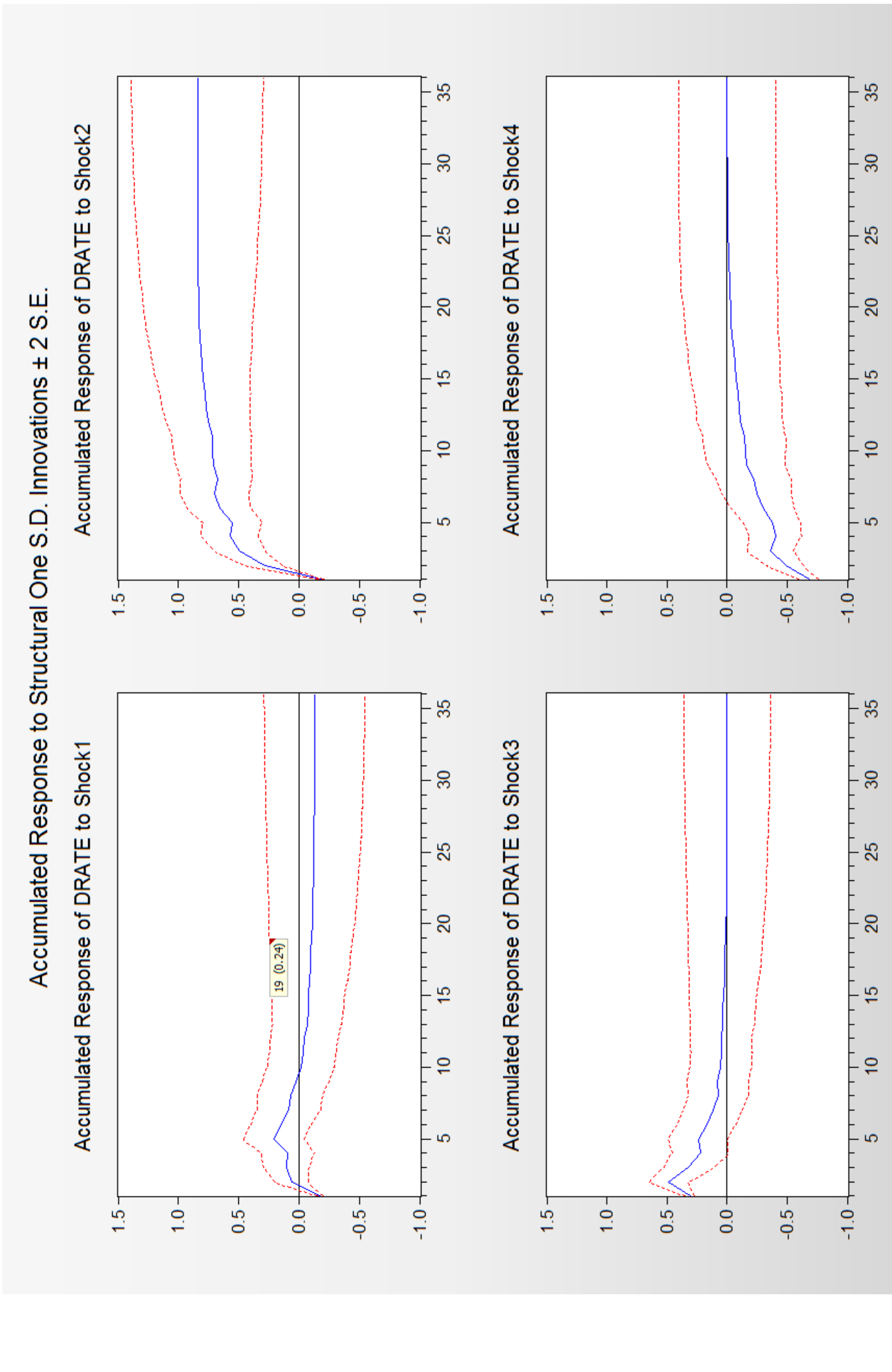


Figure 7.14: Accumulated Impulse Responses of the Interest Rate for Gali (1992) using Alternative Restrictions

Bibliography

[1] Alessi,L. M. Barigozzi and M. Capasso (2008) "A review of nonfundamentalness and identification in structural VAR models," *European Central Bank Working Paper Series 922.*

[2] An, L. (2006), "Exchange Rate Pass-Through: Evidence Based on Vector Autoregression with Sign Restrictions," manuscript, University of Kentucky.

[3] Anderson, T.W. and H. Rubin (1949), "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations", *Annals of Mathematical Statistics* 20, 46-63.

[4] Arias, J.E., Rubio-Ramírez, J.F. and Waggoner, D.F. (2014). "Inference based on SVARs identified with sign and zero restrictions: theory and applications," Dynare Working Papers 30, CEPREMAP.

[5] Bache, I and K. Leitemo (2008), "The price puzzle: Mixing the temporary and permanent monetary policy shocks." *Working Paper 2008/18*, Norges Bank.

[6] Bai, J. and S. Ng (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74, 1133–1150.

[7] Bai, J. and S. Ng (2013), "Principal Components Estimation and Identification of Static Factors," *Journal of Econometrics*, 176, 18-29.

[8] Baillie, R. T., 1996, Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics*, 73, 5–59.

[9] Bańbura, M., G. Domenico and M. Lenza, 2015. "Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections," *International Journal of Forecasting*, 31(3), 739-756.

[10] Baumeister, C. and Hamilton, J.D. (2014). "Sign restrictions, structural vector autoregressions, and useful prior information", NBER Working Papers 20741

[11] Bellégo,C. and L. Ferrara (2009), "Forecasting Euro-Area Recessions Using Time-Varying Binary Response Models for Financial Variables," *Document de Travail No. 259*, Banque de France.

[12] Berg,A. P. Karam and D. Laxton (2006), "A Practical Model-Based Approach to Monetary Policy Analysis–Overview," *IMF Working Papers , 06/80.*

[13] Bernanke, B.S., J. Bovian and P. Eliasz (2005), "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach," *Quarterly Journal of Economics* 120:387-422.

[14] Blanchard, O. and R. Perotti (2002), "An Empirical Characterization Of The Dynamic Effects Of Changes In Government Spending And Taxes On Output," *The Quarterly Journal of Economics*, 117(4), 1329-1368.

[15] Boivin, J. and M.P. Giannoni (2009), "Global Forces and Monetary Policy Effectiveness," in J. Gali and M. Gertler (eds), *International Dimensions of Monetary Policy*, University of Chicago Press, Chicago, 429-478.

[16] Bjørnland, H. and J. Halvorsen (2008), "How Does Monetary Policy Respond to Exchange Rate Movements? New International Evidence," *Norges Bank Working Paper #2008/15.*

[17] Blanchard, O. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review*, 79, 655-673.

[18] Canova, F (2007), *Methods for Applied Macroeconomic Research.* Princeton University Press.

[19] Canova, F. and G. De Nicoló (2002), "Monetary Disturbances Matter for Business Fluctuations in the G-7," *Journal of Monetary Economics*, 49, 1131-1159.

[20] Canova, F. and M. Paustian (2010), "Measurement with Some Theory: Using Sign Restrictions to Evaluate Business Cycle Models," Economics Working Papers, Department of Economics and Business, Universitat Pompeu Fabra.

[21] Canova, F. and G. De Nicoló (2002), "Monetary Disturbances Matter for Business Fluctuations in the G-7," *Journal of Monetary Economics*, 49, 1131-1159.

[22] Castelnuovo, E. and P. Surico (2010), "Monetary Policy, Inflation Expectations and the Price Puzzle," *Economic Journal*, 120, 1262-1283

[23] Catão, L and A. Pagan (2011) , "The Credit Channel and Monetary Transmission in Brazil and Chile: A Structured VAR Approach," L.F. Cespedes, R. Chang and D. Saravia (eds) *Monetary Policy Under Financial Turbulence* (Central Bank of Chile).

[24] Cherif, R. and F. Hasanov (2012), "Public Debt Dynamics: The Effects of Austerity, Inflation and Growth Shocks," *IMF Working Paper*, 12/230.

[25] Chevillon, G., S. Mavroeidis and Z. Zhan (2015) "Robust inference in structural VARs with long-run restrictions" mimeo, Oxford University

[26] Cho, S. and A. Moreno (2006), "A Small-Sample Study of the New Keynesian Macro Model," *Journal of Money, Credit and Banking*, 38, 1461-1481.

[27] Christiano, L.J., M. Eichenbaum and R. Vigfusson (2006), "Alternative Procedures for Estimating Vector Autoregressions Identified with Long-Run Restrictions," Journal of the European Economic Association, 4, 475-483.

[28] Dedola, L. and S. Neri (2006), "What Does a Technology Shock Do? A VAR Analysis with Model Based Sign Restrictions," *ECB Working Paper* #705.

[29] Del Negro, M., and F. Schorfheide, 2010. "Bayesian Macroeconometrics.", Handbook of Bayesian Econometrics.

[30] Dungey, M. and R.A. Fry, (2009), "Identifying Fiscal and Monetary Policy in a Structural VAR," *Economic Modelling*, 26, 1147-1160.

[31] Dungey, M. and A.R. Pagan (2000), "A Structural VAR Model of the Australian Economy," *Economic Record*, 76, 321-342.

[32] Durbin, J. (1954), "Errors in Variables," *Revue of the International Statistical Institute*, 22, 23-32.

[33] Evans, L. and G. Wells (1983), "An Alternative Approach to Simulating VAR Models," *Economics Letters*, 12, 23-29.

[34] Evans, L. and G. Wells (1986), "Confidence Regions for Multipliers in Linear Dynamic Models," *Econometrica*, 54, 699-706.

[35] Eickmeier, S. and B. Hofmann (2013), "Monetary Policy, Housing Booms, And Financial (Im)Balances," *Macroeconomic Dynamics*, Cambridge University Press, 17, 830-860.

[36] Farrant, K. and G. Peersman (2006), "Is the Exchange Rate a Shock Absorber or a Source of Shocks? New Empirical Evidence," *Journal of Money, Credit and Banking,* 38, 939-961.

[37] Faust, J. (1998). "The Robustness of Identified VAR Conclusions about Money," *Carnegie–Rochester Conference Series on Public Policy*, 49, 207-44.

[38] Faust, J. and E.M. Leeper (1997), "When do Long-Run Identifying Restrictions Give Reliable Results," *Journal of Business and Economic Statistics*, 15, 345-353.

[39] Favero, C. and F. Giavazzi (2007), "Debt and the Effects of Fiscal Policy*,"* *NBER Working Papers 12822*, National Bureau of Economic Research, Inc.

[40] Fisher, L.A., H-S Huh and A.R. Pagan (2015), "Econometric Methods for Modelling Systems with a Mixture of I(1) and I(0) Variables," *Journal of Applied Econometrics* (forthcoming).

[41] Francis, N., M.T. Owyang and J.E. Roush (2005), "A Flexible Finite-Horizon Identification of Technology Shocks," *Board of the Governors of the Federal Reserve System, International Finance Discussion Papers #832.*

[42] Francis, N., M.T. Owyang and A.T. Theodorou (2003), "The Use of Long-Run Restrictions for the Identification of Technology Shocks," *Federal Reserve Bank of St Louis* Review, *November-December, 53-66.*

[43] Frühwirth-Schnatter, S. (2001), "Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models," Journal of the American Statistical Association, 96, 194-209.

[44] Fry, R.A. and A.R. Pagan (2007), "Some Issues in Using Sign Restrictions for Identifying Structural VARs," *NCER Working Paper* #14.

[45] Fry, R. and A.R. Pagan,"Sign Restrictions in Structural Vector Autoregressions: A Critical Review," *Journal of Economic Literature*, 2011, 49, 938-960.

[46] Fujita, S. (2009), "Dynamics of Worker Flows and Vacancies: Evidence from the Sign Restriction Approach," *Journal of Applied Econometrics*, forthcoming.

[47] Giordani, P. (2004), "An Alternative Explanation of the Price Puzzle," *Journal of Monetary Economics*, 51, 1271-1296.

[48] Gali, J. (1992), "How Well does the ISLM Model Fit Postwar U.S. Data?" *Quarterly Journal of Economics*, 107, 709-735.

[49] Gali. J. (1999), "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *The American Economic Review*, 89, 249-271.

[50] Giratis, L, G. Kapetanios and T. Yates (2014), "Inference on Stochastic Time-Varying Coefficient Models," *Journal of Econometrics*, 179, 46-65.

[51] Gonzalez, A. (2016), "Bayesian VARs classnotes" IMF - ICD course.

[52] Gospodinov, N., 2010. "Inference in Nearly Nonstationary SVAR Models With Long-Run Identifying Restrictions," *Journal of Business and Economic Statistics*, 28, 1-12.

[53] Gordon, D.B. and E.M. Leeper (1994), "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification," *Journal of Political Economy*, 102 (1994), 1228-1247.

[54] Gouriéroux, C., and A. Monfort (2014): "Revisiting identification and estimation in structural VARMA models," *CREST Discussion Paper 2014-30.*

[55] Hansen, B. "Threshold Autoregression in Economics," *Statistics and its Interface,* 4 (2011), 123-127.

[56] Hansen, L.P. and T.J. Sargent (1991), "Two Difficulties in Interpreting Vector Autoregressions," in L.P. Hansen and T.J. Sargent (1991), *Rational Expectations Econometrics*, (Westview Press, Boulder), 77-120.

[57] Harding, D., and Pagan, A.R. (2002), "Dissecting the Cycle: A Methodological Investigation," *Journal of Monetary Economics*, 49, 365-381.

[58] Harding, D. and A.R. Pagan, (2011) "An Econometric Analysis of Some Models of Constructed Binary Random Variables," *Journal of Business and Economic Statistics*, 29, 86-95.

[59] Hau, H. and H. Rey (2004), "Can Portfolio Rebalancing Explain the Dynamics of Equity Returns, Equity Flows and Exchange Rates," *American Economic Review*, 94, 126-133.

[60] Hausman, J. A, (1975), "An Instrumental Variable Approach to Full Information Estimators for Linear and Certain Nonlinear Econometric Models," *Econometrica*, 43, 727-38.

[61] Heinlin, R. and H-M Krolzig (2011), "Effects of monetary policy on the Pound-Dollar exchange rate: Is there a 'delayed overshooting puzzle'?" paper given at the EC2 meeting, European University Institute.

[62] Hendry, D. F., and Krolzig, H.-M. (2005), "The properties of automatic GETS modelling," *Economic Journal*, 115, C62-C98.

[63] Hendry, D. F., Johansen, S., and Santos, C. (2008), "Automatic selection of indicators in a fully saturated regression," *Computational Statistics*, 33, 317-335. *Erratum*, 337-339.

[64] Hendry, D. F., and Krolzig, H.-M. (2005), "The properties of automatic GETS modelling," *Economic Journal*, 115, C62-C98.

[65] Herwartz.H. and H. Luetkepohl, 2011. "Structural Vector Autoregressions with Markov Switching: Combining Conventional with Statistical Identification of Shocks," *Economics Working Papers ECO2011/11, European University Institute.*

[66] Herwartz, H. and M. Plodt (2016), "The Macroeconomic Effects of Oil Price Shocks: Evidence from a Statistical Identification Approach", *Journal of International Money and Finance*, 61, 30-44.

[67] Iacoviello, M. (2005), "House Prices, Borrowing Constraints and Monetary Policy in the Business Cycle," *American Economic Review*, 95, 739-764.

[68] Jääskelä, J. and D. Jennings (2011), "Monetary Policy and the Exchange Rate: Evaluation of VAR Models," *Journal of International Money and Finance,* 30, 1358-1374.

[69] Jarociński, M. and F.R. Smets (2008), "House Prices and the Stance of Monetary Policy," *Review*, Federal Reserve Bank of St Louis, July/August 2008, 339-365.

[70] Justiniano, A. and B. Preston (2010), "Monetary Policy Uncertainty in an Empirical Small Open Economy Model," *Journal of Applied Econometrics*, 25, 93-128.

[71] Kadiyala, K., and S. Karlsson (1997), "Numerical methods for estimation and inference in Bayesian VAR models." *Journal of Applied Econometrics,* 12, 99-132.

[72] Kapetanios,G. A.R. Pagan and A. Scott (2007) "Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modeling," *Journal of Econometrics*, 136, 505-594.

[73] Kaufmann, S. (2010), "Dating and Forecasting Turning Points by Bayesian Clustering with Dynamic Structure: A Suggestion with an Application to Austrian Data," *Journal of Applied Econometrics*, 23, 309-344.

[74] Kilian, L. (1998), "Small-Sample Confidence Intervals For Impulse Response Functions," *The Review of Economics and Statistics*, 80, 218-230.

[75] Kilian, L. and D. Murphy (2009), "Why Agnostic Sign Restrictions are Not Enough: Understanding the Dynamics of Oil Market VAR Models," J*ournal of the European Economic Association* (2012), 10, 1166-1188.

[76] Kilian, L. (2008), "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?" *Review of Economics and Statistics*, 90, 216-240.

[77] Kim, S and N. Roubini (2000), "Exchange Rate Anomalies in the Industrial Countries : A Solution with a Structural VAR Aprpoach," *Journal of Monetary Economics*, 45, 561-586.

[78] Klein, L.R. and A.S. Goldberger (1964), *An Econometric Model of the United States 1929-1952*, North Holland Publishing Company, Amsterdam.

[79] Koop, G., and D. Korobilis, 2010. "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics." *Foundations and Trends in Econometrics,* 3(4), 267-358.

[80] Krolzig, H-M and J. Toro (2004) "Classical and modern business cycle measurement: The European case," *Spanish Economic Review*, 7, 1-21.

[81] Lanne. M. and H. Lütkepohl, (2008), "Identifying Monetary Policy Shocks via Changes in Volatility," *Journal of Money, Credit and Banking*, 40, 1131-1149.

[82] Lanne M, Meitz M,and P. Saikkonen (2015), "Identification and estimation of non-Gaussian structural vector autoregressions", CREATES Research Papers 2015-16, School of Economics and Management, University of Aarhus.

[83] Lutkepohl, H. (2006). *New Introduction to Multiple Time Series Analysis.* Springer.

[84] Lütkepohl, H. and P. Saikkonen (1997), "Impulse response analysis in infinite order cointegrated vector autoregressive processes," *Journal of Econometrics*, 81, 127-157.

[85] Leamer, E. (1984), "Vector Autoregressions for Causal Inference," *Carnegie-Rochester Conference Series on Public Policy*, 22, 255-303.

[86] Lewis, V.J. (2007), "Productivity and the Euro-Dollar Real Exchange Rate," *Review of World Economics*, 143, 324-347.

[87] Lippi, M. and L. Reichlin (1994), "VAR Analysis, Nonfundamental Representations, Blaschke Matrices," *Journal of Econometrics*, 63, 644-52.

[88] Litterman, R., (1986) "Forecasting with Bayesian vector autoregressions -five years of experience." *Journal of Business and Economic Statistics*, Vol. 4, 25-38.

[89] Lütkepohl, H. (2014), "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey," *DIW Discussion Paper 1351 , German Institute for Economic Research, Berlin.*

[90] Marcellino, M., J. Stock, and M. Watson, (2006), "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macreconomic Time Series," *Journal of Econometrics* 127, 499-526.

[91] McKibbin, W., A.R.Pagan and J. Roberston (1998), "Some Experiments in Constructing a Hybrid Model for Macroeconomic Analysis," *Carnegie-Rochester Series in Public Policy*, 49, 113-142.

[92] Mertens, K. and M.O Ravn (2012), "Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks," *American Economic Journal: Economic Policy,* 4, 145-81.

[93] Mountford, A. (2005), "Leaning into the Wind: A Structural VAR Investigation of UK Monetary Policy," *Oxford Bulletin of Economics and Statistics*, 67, 597-621.

[94] Mountford, A. and H. Uhlig (2005), "What are the Effects of Fiscal Policy Shocks," *Discussion Paper 2005-039, SFB 649*, Humboldt-Universitat, Berlin.

[95] Mountford, A. and H. Uhlig (2008), "What are the Effects of Fiscal Policy Shocks?" *NBER Working Paper* #14551.

[96] Olea, J.L.M, J.H. Stock and M.W. Watson (2013), "Inference in Structural VARs with External Instruments," paper given to the 3rd VALE-EPGE Global Economic Conference "Business Cycles".

[97] Ouliaris, S., and A. Pagan (2016), "A Method for Working With Sign Restrictions in Structural Equation Modelling", *Oxford Bulletin of Economics and Statistics* (forthcoming).

[98] Pagan, A.R. and M.H. Pesaran (2008), "Econometric Analysis of Structural Systems with Permanent and Transitory Shocks," *Journal of Economic Dynamics and Control*, 32, 3376-3395.

[99] Pagan, A.R. and J. Robertson (1998), "Structural Models of the Liquidity Effect," *Review of Economics and Statistics*, 80, 202-217.

[100] Pagan, A.R. and T. Robinson (2016), "Investigating the Relationship Between DSGE and SVAR Models" *NCER Working Paper* # 112.

[101] Preston, A.J. (1978), "Concepts of Structure and Model Identifiability for Econometric Systems," in A.R. Bergstrom *et al.*, *Stability and Inflation,* Wiley, 275-297.

[102] Paustian, M. (2007), "Assessing Sign Restrictions," *The B.E. Journal of Macroeconomics*, 7, 1-37.

[103] Peersman, G. (2005), "What Caused the Early Millennium Slowdown? Evidence Based on Autoregressions," *Journal of Applied Econometrics*, 20, 185-207.

[104] Peersman, G. and R. Straub (2006), "Putting the New Keynesian Model to a Test," *IMF Working Paper*, 06/135.

[105] Peersman, G. and R. Straub (2009), "Technology Shocks and Robust Sign Restrictions in a Euro Area SVAR," *International Economic Review*, 50, 727-750.

[106] Pesaran, M. H. and Y. Shin (1998), "Generalized Impulse Response Analysis in Linear Multivariate Models," *Economics Letters*, 58, 17-29.

[107] Preston, A.J. (1978), "Concepts of Structure and Model Identifiability for Econometric Systems," in A.R. Bergstrom *et al.*, *Stability and Inflation* (Wiley, 1978), 275-297.

[108] Quenouille, M.H. (1957), *The Analysis of Multiple Time Series,* Griffin's Statistical Monograohs and Course No 1 (Griffin, London).

[109] Rafiq, M.S. and S.K. Mallick (2008), "The Effects of Monetary Policy on Output in EMU3 A Sign Restriction Approach," *Journal of Macroeconomics*, 30, 1756-1791.

[110] Rigobon, R. (2003), "Identification Through Heteroskedasticity," *Review of Economics and Statistics*, 85, 777-792.

[111] Robertson, J., and E. Tallman (1999),. "Vector autoregression: forecasting and reality." FED of Atlanta *Economic Review*, first quarter.

[112] Romer, C.D. and D.H. Romer (2010), "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763-801.

[113] Rubio-Ramírez, J.F., D.F. Waggoner and T. Zha (2006), "Markov-Switching Structural Vector Autoregressions: Theory and Application," *Computing in Economics and Finance* 69, Society for Computational Economics.

[114] Rüffer, R., M. Sanchez and J-G Shen (2007), "Emerging Asia's Growth and Integration: How Autonomous are Business Cycles," *ECB Working Paper* #715.

[115] Sargent, T and C. Sims (1977), "Business Cycle Modeling without Pretending to Have too Much a-priori Theory," in C. Sims *et al.* (eds) *New Methods in Business Cycle Research* (Minneapolis, Federal Reserve Bank of Minneapolis).

[116] Sanchez, M. (2007), "What Drives Business Cycles and International Trade in Emerging Market Economies?" *ECB Working Paper* #730.

[117] Sargent, T.J., and C.A. Sims (1977), "Business cycle modeling without pretending to have too much a-priori economic theory," in: C. Sims *et al.*, eds., *New Methods in Business Cycle Research* (Federal Reserve Bank of Minneapolis, Minneapolis).

[118] Scholl, A. and H. Uhlig, (2008), "New Evidence on the Puzzles: Results from Agnostic Identification on Monetary Policy and Exchange Rates," *Journal of International Economics*, 76, 1-13.

[119] Shapiro M and Watson M. (1988), "Sources of Business Cycle Fluctuations," in *NBER Macroeconomics Annual* 3: 111-148.

[120] Sims, C. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1-48.

[121] Sims, C. (1993), "A nine-variable probabilistic macroeconomic forecasting model." In J. Stock and M. Watson, eds., *Business Cycles, Indicators and Forecasting.* University of Chicago Press for the NBER.

[122] Sims, C. and T. Zha (1998), "Bayesian methods for dynamic multivariate models." *International Economic Review* 39:949-68.

[123] Smith, P.A. and P.M. Summers (2004), "Identification and Normalization in Markov Switching Models of 'Business Cycles," *Research Working Paper 04-09* Federal Reserve Bank of Kansas City

[124] Stephens, M. (2000), "Dealing with Label Switching in Mixture Models," *Journal of the Royal Statistical Society*, Series B , 62,795-809.

[125] Theil, H. and A. Goldberger (1960), "On pure and mixed statistical estimation in economics." *International Economic Review* 2, 65-78.

[126] Uhlig, H. (2005), "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381-419.

[127] Tinbergen, J. (1936), "Kan hier te lande, al dan niet na overheidsingrijpen, een verbetering van de binnenlandse conjunctuur intreden, ook zonder verbetering van onze exportpositie?" in *Prae-Adviezen Van de Vereeniging voor Staathuishoudkunde en de Statistiek,* Martinus Nijhoff, 's-Gravenhage, 62-108.

[128] Tinbergen, J. (1939), *Statistical testing of business cycle theories: Business cycles in the United States of America 1919–1932,* volume II, League of Nations, Geneva.

[129] Todd, R. (1984) "Improving economic forecasting with Bayesian vector autoregressions." Minnesota FED Quarterly Review, 8(4), 18-29.

[130] Vargas-Silva, C. (2008), "Monetary Policy and US Housing: A VAR Imposing Sign Restrictions," *Journal of Macroeconomics*, 30, 977-990.

[131] Waggoner, D.F. and T. Zha, (1999) "Conditional Forecasts In Dynamic Multivariate Models," *The Review of Economics and Statistics*, 81(4), 639-665.

[132] Wallis, K.F. (1977), "Multiple Time Series and the Final Form of Econometric Models," *Econometrica*, 45, 1481-1497.

[133] West, Kenneth (2006), "Forecast Evaluation," in *Handbook of Economic Forecasting*, Elliot, G. Granger C.W.J., Timmermann, A. (eds.), North Holland.

[134] Wold, H.O.A. (1949), "Statistical Estimation of Economic relationships," *Econometrica*, 17, 9-21.

[135] Wold, H O. A. (1951), "Dynamic Systems of the Recursive Type–Economic and Statistical Aspects," *Sankhya*, 11, 205-17.

[136] Zellner, A. and F. Palm (1974), "Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, 2, 17-54.