NCER National Centre for Econometric Research

NCER Working Paper Series

Some Issues in Using Sign Restrictions for Identifying Structural VARs

Renee Fry and Adrian Pagan

Working Paper #14

April 2007

Some Issues in Using Sign Restrictions for Identifying Structural VARs *

Renée Fry†and Adrian Pagan‡

April 13, 2007

Contents

1 **Introduction** **2**

2 **Structural Impulses** **4**

3 **Sign Restriction Methods for Estimating Impact Impulses** **5**
- 3.1 The Basic Strategy 5
- 3.2 Generating Orthogonal Matrices 7
  - 3.2.1 Givens Matrices 7
  - 3.2.2 Householder Transformations 9
- 3.3 Summarizing the Information 9

4 **What Do Sign Restrictions Do? A Simple Demand/Supply Example** **12**

5 **An Empirical Example- Peersman's (2005) Study** **14**

6 **Conclusion** **16**

*We would like to thank Jim Nason and Jon Faust for comments on an earlier version of this paper.

†Centre for Applied Macroeconomic Analysis, Australian National University, Canberra, A.C.T. Australia 0200, and CERF, University of Cambridge.

‡Queensland University of Technology and University of New South Wales. Pagan's research was supported by ESRC Grant No. 000 23-0244.

7 References **19**

Abstract

The paper looks at estimation of structural VARs with sign restrictions. Since sign restrictions do not generate a unique model it is necessary to find some way of summarizing the information they yield. Existing methods present impulse responses from different models and it is argued that they should come from a common model. If this is not done the implied shocks implicit in the impulse responses will not be orthogonal. A method is described that tries to resolve this difficulty. It works with a common model whose impulse responses are as close as possible to the median values of the impulse responses (taken over the range of models satisfying the sign restrictions). Using a simple demand and supply model it is shown that there is no reason to think that sign restrictions will generate better quantitative estimates of the effects of shocks than existing methods such as assuming a system is recursive.

1 Introduction

Vector Autoregressions (VAR) have become one of the most widely used tools in macroeconometric research. Although they have many uses in data description and forecasting that do not require one to identify the shocks underlying them, once one comes to consider policy actions or to understand economic outcomes it is necessary to label the shocks. Essentially this is a question of how to convert a VAR into a Structural VAR (SVAR) and how to identify the parameters of such a system. Initially the standard way of doing this was to assume that the macroeconomic system could be represented as a set of simultaneous equations that recursively determined economic variables. Later some non-recursivity was allowed and often the requisite identifying information became inertial restrictions e.g. monetary policy had no impact on real variables for two quarters. In more recent times there has been a move away from the imposition of these short-run restrictions to either long-run restrictions or the incorporation of qualitative and quantitative information. The latter generally involves the use of either sign restrictions or other prior information upon the impulse responses.

In this paper we look at the use of sign restrictions as identifying information. Basic to this method of shock identification (and all SVARs) is the requirement that shocks be uncorrelated. Starting with a base set of shocks that satisfy this assumption, one then combines these together to create new shocks whose impulse responses are tested to see if they satisfy the postulated signs. The weights used in the construction of these impulse responses depend on a vector of parameters θ. By assigning a value to θ we produce a *model,* with associated shocks and impulse responses. In a similar fashion to the non-parametrics literature which indexes alternative models with a set of (infinite-dimensional) parameters, here a large range of models is produced by varying the values for θ.

There is considerable diversity in the literature over the number of shocks to be identified and the types of questions one wants to address with the identified impulses. One strand has been concerned with identifying a single shock. Uhlig (2005) and Faust (1998) are examples of this, where monetary policy shocks are the focus of attention, although Faust uses quantitative as well as sign restrictions. Qualitative questions are then often asked about the impulses of some specified variables to the identified shock (these impulses of course not being used to identify it) e.g. do technology shocks have a positive or negative effect upon hours? Then the central issue that immediately arises is how one is to provide an answer to this question, as there are many values of θ (models) and they may not all produce the same sign for (say) the effects of technology. One possibility is to look at what fraction of the models produce a particular sign. Quantitative information about the impulse responses may also be needed. Indeed, if one wished to use the impulse information for policy analysis, it is required. Given that the models may produce very different magnitudes for the impulse responses one needs to decide on some strategy for presenting the range of values of the impulse responses coming from the models, and what would be a representative value of the magnitude of this effect. There are various suggestions as to how this should be done and we later examine the difficulties with the most popular one.

If many shocks are identified, as in Peersman (2005), Peersman and Straub (2006), Canova and de Nicolo(2002), Scholl and Uhlig (2005), Mountford and Uhlig (2005) and Rubio-Ramirez et al. (2005), the issues noted above regarding θ are intensified, and we show that, unless one is careful, one may be presenting information about shocks that come from different models. Whether this is important or not depends upon what we want to do with the information. If one simply wants a description of the range of outcomes that are possible there is probably little to object to. But for operational use it doesn't make much sense to use the impulse responses of (say) money shocks from one model along with (say) the technology shocks from another. In order to use this information for any policy analysis one needs to have the associated impulses from a common model (to emphasize the point, it would make no sense to use a technology shock from an RBC model and a monetary policy shock from a backward looking IS-LM model -

we need a common model that incorporates both of these shocks). Another way of describing the difficulty is to recognize that, unless we use as shocks those that are associated with a single model, there is no guarantee that these shocks will be uncorrelated. Many of the uses made of the information, such as the construction of variance decompositions, require that shocks be uncorrelated. We show that existing studies do not seem to be aware of this difficulty and that it can have a substantial influence upon the conclusions about the magnitude of impulse responses. There is no unique solution to the problem but we present one that has the same "flavor" as the most popular existing solution but which ensures that one is working with a consistent model.

Our paper starts by setting out in a simple way the approach in the literature. We then construct a demand and supply model and use it to illustrate how the sign restriction methods produce a range of models distinguished by different values of the parameters of the demand and supply curves. Except in one specific instance the range of models produced by the methods does not include the true model. This example illustrates a simple point. There is no reason to think that sign restrictions are superior to other methods of identifying SVARs (such as recursiveness) once we ask about the magnitudes of impulse responses, and almost all studies in this vein present magnitudes. It may be that sign restrictions produce values for the responses that are closer to the true values, although the fact that they are compatible with many models, and each of these models is likely to have impulse responses of differing magnitudes, means that we have to decide on which value one should present. This is the same problem as was noted above.

Section 5 examines a recent paper by Peersman (2005), showing that the impulse responses of the shocks that he presents come from different models. If we require them to come from the same model the magnitudes and even signs can be quite different to what he presents. Moreover, since the variance and tracking decompositions which he provides assume that the shocks are uncorrelated, these must be invalid given that the constructed shocks do not have that feature. Other studies employing variance decompositions have potentially the same problem.

2 Structural Impulses

For simplicity we will consider a structural VAR, SVAR(1), of first order,

$$B_0 z_t = B_1 z_{t-1} + \varepsilon_t$$

with underlying VAR

$$z_t = A_1 z_{t-1} + v_t.$$

We immediately have that $v_t = B_0^{-1}\varepsilon_t$. The solution to the VAR(1) is the MA form

$$z_t = D(L)v_t$$

where $D(L) = I + D_1 L + D_2 L^2 + ...$, with the D_j being the impulse responses of z_{t+j} to a unit change in v_t. It follows that the MA form for the SVAR is

$$z_t = C(L)\varepsilon_t$$

where $C(L) = C_0 + C_1 L + ...$, with the impulse responses to ε_t being $C_j = D_j B_0^{-1} = D_j C_0$.

The VAR is easily estimated by OLS regardless of the nature of the SVAR. Hence D_j can always be found once the lag length of the VAR is specified. We assume that the VAR is of finite order as that is maintained in the sign restriction literature. Note that it is only necessary to determine C_0 using sign information as D_j can be found without it and so the only unknown in C_j is C_0. Hence we will focus upon the estimation of C_0. This explains why our simple example later has no dynamics and is concerned with identifying C_0. The same issues come up in that simple context as will come up in the more complex one where dynamics are involved.

3 Sign Restriction Methods for Estimating Impact Impulses

3.1 The Basic Strategy

Suppose that we have ordered the variables appearing in a VAR in some recursive way and then computed estimates of B_0. This will mean that the VAR residuals $\hat{v}_t$ are related to the structural residuals as $\hat{v}_t = \hat{B}_0^{-1}\hat{\varepsilon}_t$. If we call S the matrix that has the estimated standard deviations of the ε on the diagonal and zeros elsewhere, we could write $\hat{v}_t = \hat{B}_0^{-1}SS^{-1}\hat{\varepsilon}_t = T\eta_t$, where $\eta_t = S^{-1}\hat{\varepsilon}_t$ has unit variances.

Now suppose we could find a square matrix Q such that $Q'Q = QQ' = I$. Then

$$\begin{aligned}\hat{v}_t &= TQ'Q\eta_t \\ &= T^*\eta_t^*\end{aligned}$$

and we have a new set of estimated shocks η_t^* that also have the property that their covariance matrix is I since $E(\eta_t^*\eta_t^{*\prime}) = QE(\eta_t\eta_t')Q' = I$. Thus we have found a combination of the shocks η_t^* that have the same covariance matrix as η_t (and which will reproduce the $var(z_t)$) but which will have a different impact upon v_t and, hence, the variables z_t. It is this ability to create a large number of candidate shocks that is the basis of sign restriction methods.

Now the above constructs a set of shocks that are uncorrelated. Suppose we only wanted to extract a single shock, for example let us call it a monetary shock, that was uncorrelated with the remaining shocks. To do this we could just use one of the orthogonal shocks identified above. But we might want the non-monetary shocks to have a covariance matrix that is not the identity matrix i.e. they might be correlated with each other but uncorrelated with the monetary shock.

Let us call the first shock the monetary shock. Then $\eta_t^* = Q\eta_t$ and

$$Q = \begin{bmatrix} a & b \\ c & B \end{bmatrix}$$

where we have partitioned Q according to the monetary and remaining shocks. Now construct new shocks $\xi_t = Q^*\eta_t$ where

$$Q^* = \begin{bmatrix} a & b \\ c^* & B^* \end{bmatrix}$$

with $c^* = Dc, B^* = DB$ and D being non-singular. Clearly $\xi_{1t} = \eta_{1t}^*$ and this monetary shock is orthogonal to the other shocks in ξ_t. However, the latter need not be orthogonal to one another as they have covariance matrix DD'.

To find the impact of the monetary shock we note that $v_t = T\eta_t$, where T is a triangular matrix. Hence

$$\begin{aligned} v_t &= TQ^{*-1}Q^*\eta_t \\ &= TQ^{*-1}\xi_t. \end{aligned}$$

Now, since $Q^*Q^{*\prime} = I$, we have $Q^{*-1} = Q^{*\prime}$ and

$$Q^{*-1} = \begin{bmatrix} a & c^{*\prime} \\ b' & B^{*\prime} \end{bmatrix}.$$

Hence the responses are

$$\begin{aligned} TQ^{*-1} &= \begin{bmatrix} t_{11} & 0 \\ \overline{t} & \overline{T} \end{bmatrix} \begin{bmatrix} a & c^{*\prime} \\ b' & B^{*\prime} \end{bmatrix} \\ &= \begin{bmatrix} t_{11}a & t_{11}c^{*\prime} \\ a\overline{t} + \overline{T}b' & \overline{t}c^{*\prime} + \overline{T}B^{*\prime} \end{bmatrix} \end{aligned}$$

and $\begin{bmatrix} t_{11}a \\ a\bar{t} + \overline{T}b' \end{bmatrix}$ is the impact impulse responses to the monetary shock. But this is identical to what we would get using the η_t^* i.e. treating all the shocks as if they were orthogonal.

3.2 Generating Orthogonal Matrices

How do we get Q? One example of Q is simply that which involves re-ordering of the variables i.e. we maintain a recursive model but vary the nature of the recursivity. This produces a new set of shocks and impulses but the shocks will still be orthogonal. One often sees the comment that other orderings were tried with the same result in terms of impulse responses. It's also often the case that people try different orderings but then choose between them based on the signs and magnitudes of the estimated impulse responses, as one can't choose between them from the data, given that they have identical VARs. But this raises the question of whether all of the different orderings of the variables exhausts the ways of combining together the shocks while keeping them orthogonal to one another i.e. retaining the identity matrix as covariance matrix. The answer is no, and it is this fact that the sign restriction literature uses.

There are currently two ways of generating orthogonal matrices. The first of these employs the Givens transform and the second is based on Householder transforms (specifically the QR decomposition of a matrix, which is done by a sequence of Householder transforms).

3.2.1 Givens Matrices

In the context of a 4 variable VAR (the example we use later) a 4×4 Givens matrix Q_{23} has the form

$$Q_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

i.e. the matrix is the identity matrix in which the $(2,3)$ and $(3,2)$ elements have been replaced by the cosine and sine terms and θ lies between 0 and

$\pi/2$. Q_{23} is called a Givens rotation. Then

$$\begin{aligned}
Q_{23}'Q_{23} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta & 0 \\ 0 & -\cos\theta\sin\theta + \sin\theta\cos\theta & \cos^2\theta + \sin^2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= I_4 = Q_{23}Q_{23}'.
\end{aligned}$$

For a four variable system one could choose other Givens rotations Q_{12}, Q_{13}, Q_{14}, Q_{23}, Q_{24}, Q_{34} as potential combining matrices. In practice most users of the approach have used the multiple of the basic set of Givens matrices e.g. in the four variable case

$$\begin{aligned}
Q_G &= Q_{12}(\theta_1) \times Q_{13}(\theta_2) \times Q_{14}(\theta_3) \times Q_{23}(\theta_4) \\
&\quad \times Q_{24}(\theta_5) \times Q_{34}(\theta_6).
\end{aligned}$$

It's clear that Q_G is orthogonal so that shocks formed as $\eta_t^* = Q_G\eta_t$ will be orthogonal and their impact upon z_t will be $T^* = TQ_G'$.

Now, the matrix Q_G above depends upon six values for θ and, as we change θ_j, we will get different values for it. Canova and de Nicolo (2002) suggested that one make a grid of M values for each of the values of θ_j between 0 and π, and then compute the 12^6 possible Q_G. Of course all of these models distinguished by θ are observationally equivalent in that they produce an exact fit to the first two moments of the data. They therefore propose that one use sign information about the impulses to decide which of these is ineligible e.g. it might be asserted that a positive interest shock should have a negative effect upon output and inflation from 2-6 lags out, and some values of θ will produce impulse responses that fail to observe these sign restrictions. Then only those combinations that produced a shock that had such a feature would be retained for further analysis.

In recent times a quasi-Bayesian approach has become popular instead of the grid method. The θ_j are taken to be uniformly distributed over $(0, \pi)$ and then (in the four variable case) realizations are made from the product of six independent $U(0, \pi)$ densities. This is really just a useful scheme for generating values of θ_j that can be used to construct candidate Q_G matrices, rather than a Bayesian analysis per se. Uhlig (2005) and Peersman (2005) use this approach, although a Bayesian treatment is also given of the VAR coefficients which produce the base set of impulse responses that are combined together with Q_G. For any realization of the VAR coefficients and error variance one can perform the Givens rotation analysis and tabulate those that satisfy the sign restrictions.

3.2.2 Householder Transformations

The alternative method of forming an orthogonal matrix Q is to generate some random variables W from an $N(0, I_4)$ density (for a four variable VAR) and then decompose $W = Q_R R$, where Q_R is an orthogonal matrix and R is a triangular matrix. Householder transformations of a matrix are used to decompose W. The algorithm producing Q_R is often called a QR decomposition. Clearly $Q_R = I$ corresponds to the matrix used in recursive orderings. Since many draws of W can be made, one can find many Q_R. Rubio-Ramirez (2005) et al. seem to have been the first to propose this, and they have argued that, as the size of the VAR grows, this is a computationally efficient strategy relative to the Givens approach.

Is there a difference between Q_R and Q_G? In the 2 dimensional case Q_R has the form

$$Q_R = \begin{bmatrix} d & -e \\ e & d \end{bmatrix}$$

(of course either the rows could be permuted or the negative of the matrix taken and these will all produce orthogonal matrices as well), where $d < 1$ and $d^2 + e^2 = 1$. Now solving $d = \cos(\theta)$ will imply a value for θ. Then, since

$$d^2 + e^2 = \cos^2\theta + \sin^2\theta = 1$$

it must be that $e = \sin(\theta)$. Thus, in the two variable case, the Givens matrix and that found using Q_R from a QR decomposition must give identical Q matrices and, hence, shocks. In the more general case we know that Q_R is orthogonal, and this produces $n(n+1)/2$ restrictions upon the n^2 elements i.e. there are just $n(n-1)/2$ free parameters in Q_R. This is the same number of free parameters as in Q_G. In the case of $n = 4$, the number of free parameters is six. So Q_G and Q_R have the same number of free elements, suggesting that one can be converted to the other, although this seems to be much harder to prove than in the two dimensional case. One obstacle seems to be how the Givens matrices are combined into a single Q_G.

3.3 Summarizing the Information

Now one might expect that there will be more than one model that have impulse responses that will satisfy the sign restrictions. Each may be quantitatively different and so one has to ask how to present a "consensus" view of the magnitudes of the responses. A common suggestion is to compute all the impulse responses $C_j^{(k)}$ that satisfy the sign restrictions, where k indexes the different values of θ, and to then report some summary measure of these, such as a median over k of $C_j^{(k)}$. Often quantiles of the $C_j^{(k)}$ are also given as well, in order to provide an impression of the range. An alternative, used for example by Faust (1998) and Uhlig(2005), is to supplement the analysis above by choosing θ such that it minimizes a function of $C_j^{(k)}$. We have little to say about the latter strategy as it clearly revolves around whether the criterion function is a reasonable one. However, it does mean that more than sign information is being used in obtaining estimates.

Because the strategy above is to work with the quantiles of the $C_j^{(k)}$ it may seem as if it is emulating the approach when one presents quantiles of a distribution from either a Bayesian or bootstrap experiment. But it is important to recognize that the distribution here is *across models*. It has nothing to do with sampling uncertainty. Even if A and V were known with certainty there will be a question of how one proceeds whenever there are many θ. There is of course a greater range when one accounts for the uncertainty in A and V, but it does not help to understand later issues by confusing these two sources of variation.

So let us turn to whether the median of the impulse responses does provide a useful measure.[1] Let us look at the impact impulse responses. If there is just a single shock and two variables there will be a range of values for the two impulse responses $\tau_1^{(k)}, \tau_2^{(k)}$, where k indexes the values of θ (θ is a scalar here). What is often presented then as the summary measure is $med(\tau_1^{(k)})$ and $med(\tau_2^{(k)})$. Now, if $\tau_j^{(k)}$ were monotonic in θ, this would be $\tau_j(med(\theta^{(k)}))$, and so the median of the impulse responses would come from the same model, that represented by $med(\theta^{(k)})$. Since there is no guarantee of monontonicity the median impulses will generally be associated with two different values of θ i.e. come from two different models. Hence what we are generally viewing are impulse responses that cannot be simultaneously generated by a single model. This does not seem to make any sense, unless one is simply asking whether there exists some model that will generate shocks that produce those impulse responses by themselves rather than jointly.

Our solution to this problem is to choose that value of $\theta^{(k)}$ that produces impulses that are as close to the median responses as possible. This would seem to preserve the consensus view that the median is a good way of summarizing the results. To devise a criterion to do this we need to recognize that the impulses are not unit free, so that we first standardize them by sub-

[1] Although we focus on the median here the points apply to any quantile.

tracting off their median, and then divide that quantity by their standard deviation over whatever set of models we have generated that satisfy the sign restrictions. These standardized impulses are then grouped into a vector $\phi^{(k)}$ (in a two variable case ϕ is 4×1 as there are four impulses) for each value $\theta^{(k)}$. Subsequently we choose the k that minimizes $\phi^{(k)\prime}\phi^{(k)}$, and then use that $\theta^{(k)}$ to calculate impulses. Whether this strategy produces a unique k is an empirical question, although in applications we have made it turns out to do so.

The likelihood of presenting information that looks as if it comes from a single model but spans many models rises with the number of shocks and variables. To illustrate this we later work with Peersman's (2005) paper which has four variables and four shocks. θ is now a six dimensional vector and 1000 models that produced impulse responses satisfying the sign restrictions were retained for analysis . There are four sets of values for θ that correspond to the median impulse responses of *output* to the four shocks. These are given in Table 1.

Table 1 Values of θ Producing Median Impulses of Output to Shocks

	oil	supply	demand	monetary
θ_1	1.410	1.575	.176	2.409
θ_2	2.363	.408	2.197	1.837
θ_3	.252	1.456	1.675	.270
θ_4	1.588	2.152	.816	2.776
θ_5	.196	2.031	2.608	2.768
θ_6	2.922	2.415	1.618	2.158

These are very different vectors so one might expect that there will be differences between the impulse responses constrained to come from a single model (one set of values for θ) and those that come from different models. In the next section we see that this is so.

Each of these four sets can be used to construct an implied shock by selecting the appropriate column from the Q_G matrices that corresponds to them. Letting this matrix be Ψ, then the covariance matrix of the shocks that would be implied by using the median impulse responses would be $\Psi\Psi'$. Below we convert this to a correlation matrix, but with the variances left on the diagonals.

$$\begin{bmatrix} 1.694 & 0.196 & -0.120 & -0.024 \\ 0.196 & 0.448 & -0.234 & 0.092 \\ -0.120 & -0.234 & 0.865 & 0.056 \\ -0.024 & 0.092 & 0.056 & 0.993 \end{bmatrix}.$$

As evident from this matrix the implied shocks are not orthogonal and their variances are not unity. There are many implications of this fact. Apart from the fact that the assumption of uncorrelated shocks was central to the philosophy of sign restrictions, it is the case that, in the event that the correlations are non-zero, no variance decomposition exists. This problem remains even if there is only a single shock, as in Uhlig (2005), since there is nothing that ensures it is uncorrelated with the remaining shocks in the model whenever we take the impulse responses from different models.

4 What Do Sign Restrictions Do? A Simple Demand/Supply Example

It is useful to consider a simple example and see how sign restrictions produce identifying information. We take the case of a demand and a supply function with associated shocks but abstract from any dynamics i.e. the system is

$$\begin{aligned} q_t &= -\beta p_t + \varepsilon_{Dt} \\ q_t &= \gamma p_t + \varepsilon_{St} \end{aligned}$$

where $\beta > 0, \gamma > 0$ and the shocks are distinguished as demand (D) and supply (S). We have

$$p_t = \frac{\varepsilon_{Dt}}{\beta + \gamma} - \frac{\varepsilon_{St}}{\beta + \gamma}$$

and so a demand shock raises prices and a supply shock reduces them. This is the sign restriction information.

Since we need to begin with a base set of orthogonal shocks a recursive system is assumed with p_t being ordered before q_t i.e. the base system is

$$\begin{aligned} p_t &= \eta_{1t} \\ q_t + \tau p_t &= \eta_{2t}. \end{aligned}$$

The matrix of impulse responses is $\begin{bmatrix} 1 & 0 \\ -\frac{1}{\tau} & \frac{1}{\tau} \end{bmatrix}$ and the orthogonal shocks η_{1t} and η_{2t} will have different sign patterns thereby making one of them demand

and the other supply. Other possible orthogonal shocks will therefore be constructed using orthogonal matrices applied to p_t and $q_t + \tau p_t$ as a base. To analyse these we first need to look at the determinants of τ. We do this through the plim of $\hat{\tau}$ found by regressing q_t on $-p_t$, since this is the appropriate estimator given the recursive system assumption. The probability limit is found as

$$
\begin{aligned}
p\lim \hat{\tau} &= -(E(p_t^2))^{-1}E(p_t q_t) \\
&= \beta - \frac{\sigma_D^2}{\sigma_D^2 + \sigma_S^2}(\beta + \gamma) \\
&= \beta - \delta(\beta + \gamma).
\end{aligned}
$$

Now new orthogonal shocks η_{1t}^* and η_{2t}^* will be constructed using a Givens rotation as the orthonormal matrix.[2] Since there is only one Givens matrix in this case - $Q_{12}-$ we have

$$
\begin{bmatrix} \eta_{1t}^* \\ \eta_{2t}^* \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}
$$

and so

$$
\begin{aligned}
\begin{bmatrix} \eta_{1t}^* \\ \eta_{2t}^* \end{bmatrix} &= \begin{bmatrix} \eta_{1t}\cos\theta - \eta_{2t}\sin\theta \\ \eta_{1t}\sin\theta + \eta_{2t}\cos\theta \end{bmatrix} \\
&= \begin{bmatrix} p_t\cos\theta - (q_t + \tau p_t)\sin\theta \\ p_t\sin\theta + (q_t + \tau p_t)\cos\theta \end{bmatrix}.
\end{aligned}
$$

Letting $\phi_1 = \cos\theta$ and $\phi_2 = \sin\theta$ the impulses will be

$$
\begin{bmatrix} \phi_1 - \tau\phi_2 & -\phi_2 \\ \phi_2 + \phi_1\tau & \phi_1 \end{bmatrix}^{-1} = \begin{bmatrix} \phi_1 & \phi_2 \\ -(\phi_2 + \phi_1\tau) & \phi_1 - \tau\phi_2 \end{bmatrix}.
$$

Now these may not have a different sign pattern - it will depend upon the value of τ and $\theta-$ but it is equally clear that there will be values of θ that will produce opposite sign patterns. So we will have a number of candidate impulse responses indexed by θ.

The new impulse responses also define new demand and supply curves of the form

$$
\begin{aligned}
q_t &= -\frac{(\phi_2 + \tau\phi_1)}{\phi_1}p_t + (\eta_{2t}^*/\phi_1) \\
q_t &= \frac{(\phi_1 - \tau\phi_2)}{\phi_2}p_t + (\eta_{1t}^*/\phi_2).
\end{aligned}
$$

[2] As we have pointed out the same conclusion will be reached using the QR approach as there is a single parameter to be determined in that matrix, but in some ways it is easier to analytically work with the Givens matrix.

Given τ (which can be consistently estimated) we see that the demand and supply curves are made functions of a single parameter θ and this constrains the values they can take. It is natural then to ask whether we could ever expect to get correct values for the demand and supply parameters by applying sign restrictions. For that to happen we would need

$$\begin{aligned}\beta &= \frac{(\phi_2 + \tau\phi_1)}{\phi_1} = \frac{\phi_2}{\phi_1} + \tau \\ \gamma &= \frac{\phi_1}{\phi_2} - \tau\end{aligned}$$

which imply that

$$\begin{aligned}\gamma + \beta &= \frac{1}{\phi_1\phi_2} \qquad (1) \\ \gamma - \beta &= \frac{\phi_1^2 - \phi_2^2}{\phi_1\phi_2} - 2(\beta - \delta(\beta + \gamma)).\end{aligned}$$

Adding 2β to each side of the second expression and re-arranging it produces

$$\begin{aligned}(1 - 2\delta)(\gamma + \beta) &= \frac{\phi_1^2 - \phi_2^2}{\phi_1\phi_2} \qquad (2) \\ &\implies (1 - 2\delta) = \phi_1^2 - \phi_2^2. \qquad (3)\end{aligned}$$

Since (1) will determine a value for θ it will be impossible to satisfy (2) for arbitrary δ. Only in the case that $\gamma = \beta = 1, \sigma_D^2 = \sigma_S^2 = 1$ will the conditions be simultaneously satisfied since then $\delta = \frac{1}{2}$, $\theta = \frac{\pi}{4}$ from (2) and, for this θ, $\frac{1}{\phi_1\phi_2} = 2$.

It should probably not be surprising that one cannot recover the correct elasticities simply by the use of sign restrictions, since sign restrictions are very weak information. But the literature largely treats them as if they are capable of recovering accurate quantitative information. What this example shows is that there is no reason to suppose that sign restrictions are better than any other way of eliciting information on impulse responses, such as provided by short-run or long-run restrictions.

5 An Empirical Example- Peersman's (2005) Study

Peersman (2005) estimates a four-variate VAR for the Euro region and the US. The variables are the first difference of the log of oil prices (Δoil_t), output growth (Δy_t), consumer price inflation (Δp_t), and the short term nominal interest rate (s_t). In this section we replicate the model only for the US data, which involves a quarterly VAR(3) estimated over the period 1980Q1 to 2002Q2, with both a constant and a time trend included. Peersman identifies four shocks using the sign restriction methodology. These are a demand shock $\left(\varepsilon_t^D\right)$, a monetary policy shock $\left(\varepsilon_t^M\right)$, and two supply shocks: the first being an oil price shock $\left(\varepsilon_t^O\right)$, while the second is labelled a supply shock $\left(\varepsilon_t^S\right)$. The method of generating candidate impulse responses is the same as described earlier.

Sign restrictions are only applied to the contemporaneous effects of oil prices and interest rate shocks but, for the other two shocks, the sign of impulses over four periods is used. Specifically, a positive demand shock is expected to generate a positive response in oil prices (oil_t), output (y_t), prices (p_t), and the interest rate as follows:

$$
\begin{aligned}
C_{oil,j}^{\varepsilon D} &\geq 0, \; j = 0, \\
C_{y,j}^{\varepsilon D} &\geq 0, \; j = 0, 1, 2, 3, \\
C_{p,j}^{\varepsilon D} &\geq 0, \; j = 0, 1, 2, 3, \\
C_{s,j}^{\varepsilon D} &\geq 0, \; j = 0.
\end{aligned}
$$

A positive monetary policy shock is expected to generate a negative response in oil prices, output and the price level, and a positive response in interest rates. This leads to the restrictions:

$$
\begin{aligned}
C_{oil,j}^{\varepsilon M} &\leq 0, \; j = 0, \\
C_{y,j}^{\varepsilon M} &\leq 0, \; j = 0, 1, 2, 3, \\
C_{p,j}^{\varepsilon M} &\leq 0, \; j = 0, 1, 2, 3, \\
C_{s,j}^{\varepsilon M} &\geq 0, \; j = 0.
\end{aligned}
$$

A favourable supply shock is expected to lead to an increase in output and a decline in prices and the interest rate. A positive oil price shock has a positive impact on oil prices, while the effect of a supply shock on oil is ambiguous. To distinguish between the oil shock and the supply shock, it is assumed that the impact of an oil price shock on oil prices is larger than a supply shock on oil prices. Thus a positive supply shock has these effects:

$$
\begin{aligned}
C_{y,j}^{\varepsilon S} &\geq 0, \; j = 0, 1, 2, 3, \\
C_{p,j}^{\varepsilon S} &\leq 0, \; j = 0, 1, 2, 3, \\
C_{s,j}^{\varepsilon S} &\geq 0, \; j = 0.
\end{aligned}
$$

Finally, a positive oil price shock involves

$$\begin{aligned} C^{\varepsilon O}_{oil,j} &\geq 0,\ j = 0, \\ C^{\varepsilon O}_{y,j} &\leq 0,\ j = 0, 1, 2, 3, \\ C^{\varepsilon O}_{p,j} &\geq 0,\ j = 0, 1, 2, 3, \\ C^{\varepsilon O}_{s,j} &\geq 0,\ j = 0, \end{aligned}$$

and

$$C^{\varepsilon O}_{oil,j} \geq C^{\varepsilon S}_{oil,.j},\ j = 0.$$

In line with our objective of focussing attention upon the problems of dealing with multiple models we did not follow his Bayesian methodology to estimate the parameters of the VAR. Rather, the VAR was estimated using OLS and then these coefficients were treated as fixed. Then θ were drawn as described earlier. On average, 109 values of θ have to be drawn to generate a single set of impulses which satisfy all restrictions. In line with his analysis Figure 1 presents the medians of the impulse responses. As we have already commented, these do not come from a single model. We therefore applied our alternative method in which the impulse responses come from a single model chosen so that its responses are as close as possible to the median values taken over all the models satisying the sign restrictions. These are then given by the dashed lines in Figure 1. It is clear that there are sometimes very different quantitative values assigned to the impulse responses e.g. the effects of monetary shocks upon output and prices.

Figure 2 shows a decomposition of output growth into the contribution from various shocks. Peersman's decomposition is based on the medians of the contributions of each shock that come from each model whereas our decomposition uses the impulse responses and shocks that come from the single model which stems from our close-to-median rule. There is clearly much more impact of shocks in the single model case. The contribution of shocks over the 1995-2002 period to the sum of squared output growth in Peersman's case is just 9.6% compared to the 72.2% that comes from imposing a single model.

6 Conclusion

It is a powerful adage that weak information produces weak results, and sign restrictions are weak information. It is extraordinary how we constantly think that this adage only applies to past methods and that, somehow, the new method will be exempt from it. Based on the frequency of use, and

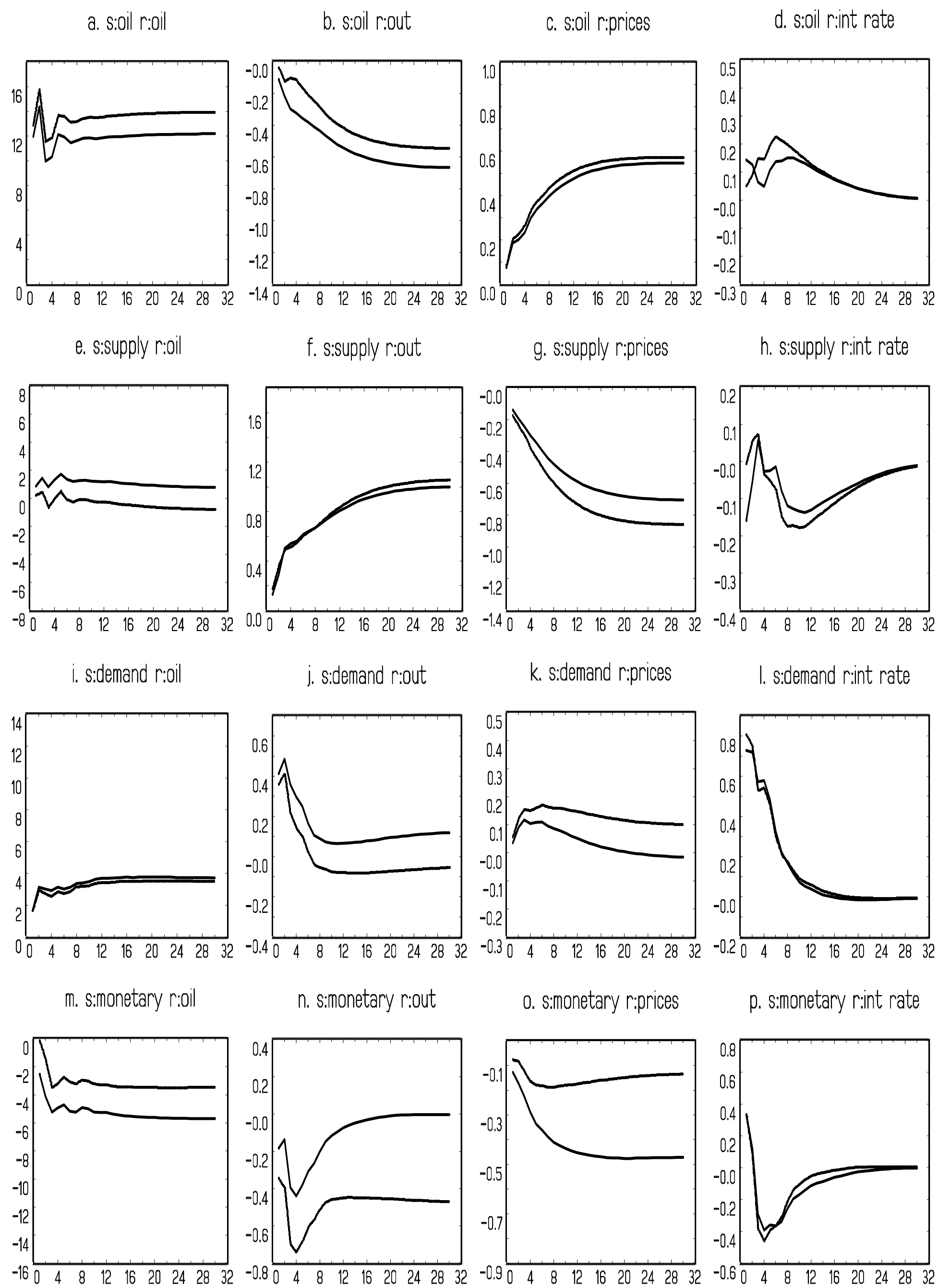

Figure 1: Peersman (2005) median $med\left(C_j^{(k)}\right)$ impulse responses and those from the suggested rule.

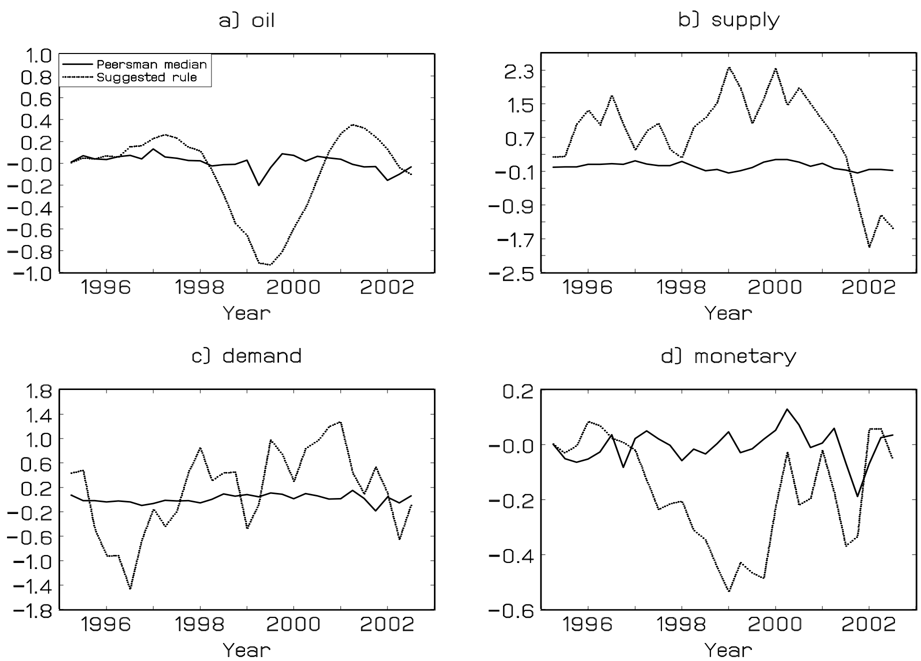

Figure 2: Historical decompositions 1995Q1-2002Q2. Contribution of shocks to output growth, comparison of Peerman and suggested rule.

some of the claims made about this method, one would easily be led to think that sign restrictions will provide better estimates of impulse responses than other existing methods. We hope to have shown that this is far from the truth. This does not mean that this method cannot be useful. But there are problems in how it is currently being used. The biggest problem is that the sign restriction methodology needs to deal with is the fact that a multiplicity of models is generated by the analysis. The most popular method currently in use presents impulse responses from different models rather than a single model and this would seem to be misleading as a description of a single economy. Moreover this difficulty has been accentuated by the fact that techniques like variance decompositions are used, as these require that the impulse responses be uncorrelated i.e. come from a single model. We have shown an example where it makes a difference quantitatively and have made a suggestion about how to deal with that problem. No doubt there may be better methods of doing this.

7 References

Canova, F. and G. de Nicolo (2002), "Monetary Disturbances Matter for Business Fluctuations in the G7", *Journal of Monetary Economics*, 49, 1131-1159.

Faust, J. (1998), "The Robustness of Identified VAR Conclusions about Money", *Carnegie-Rochester Series on Public Policy*, 49, 207-244.

Mountford, A. and H. Uhlig (2005), "What are the Effects of Fiscal Policy Shocks", *Discussion Paper 2005-039, SFB 649, Humboldt-Universitat, Berlin.*

Peersman, G. (2005), "What Caused the Early Millenium Slowdown? Evidence Based on Autoregressions", *Journal of Applied Econometrics*, 20, 185-207.

Peersman, G. and R. Straub (2006), "Putting the New Keynesian Model to a Test", *IMF Working Paper*, 06/135.

Rubio-Ramirez, J.F., D.F. Waggoner and T. Zha (2005), "Markov-Switching Structural Vector Autoregressions: Theory and Application", *Working Paper 2005-27, Federal Reserve Bank of Atlanta.*

Scholl, A. and H. Uhlig (2005), "New Evidence on the Puzzles: Results from Agnostic Indetification of Monetary Policy and Exchange Rates", *Discussion Paper 2005-037, SFB 649, Humboldt-Universitat, Berlin.*

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.