Structural VARs II

May 10, 2012

Structural VARs

Today:

- Long-run restrictions
- Two critiques of SVARs

Blanchard and Quah (1989), Rudebusch (1998), Gali (1999) and Chari, Kehoe McGrattan (2008).

Recap:

We know how to estimate VAR(p) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t : \varepsilon_t \sim N(0, \Omega)$$

But sometimes we are interested in the *structural form* of a VAR

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \quad (1)$$

Recap:

Last time we discussed how to estimate A_0 using the Choleski decomposition

- This implied ordering the variables according to contemporaneous causality

Whether this is a good idea or not cannot be judged by simply looking at the data

- Identifying assumptions need to be motivated carefully and appear sensible on an a priori basis

Recap:

To go from *reduced form*

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \\ &= A_0^{-1} A_1 y_{t-1} + A_0^{-1} A_2 y_{t-2} + ... + A_0^{-1} A_p y_{t-p} + A_0^{-1} \mathbf{u}_t \end{aligned}$$

to the *structural form* we assumed that A_0 is lower triangular

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \mathbf{u}_t$$

so that A_0^{-1} =chol(Ω).

- Important: There are (infinitely) many ways to orthogonalize the reduced form errors ε_t that all fit the data equally well:
  - There exists an infinite number of orthonormal matrices T such that $TT' = I$ and hence $A_0^{-1} TT' \left(A_0^{-1}\right)' = \Omega$

Long run restrictions

Another class of “weak” restriction that may hold across a large range of models are so called long run restrictions

- First suggested by Blanchard and Quah (1989)

The idea: There may a be a subset of shocks that have permanent effects on some variables but not on others, and shocks that have no permanent effects on any variables

- This splits reduced form shocks into permanent shocks and "everything else"

Blanchard and Quah's model of unemployment and GNP growth

A simplified version (e.g. VAR(1) and demeaned data) can be written as

$$\begin{bmatrix} a_{11}^0 & a_{12}^0 \\ a_{21}^0 & a_{22}^0 \end{bmatrix} \begin{bmatrix} U_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix} \begin{bmatrix} U_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \begin{bmatrix} u_t^d \\ u_t^s \end{bmatrix}$$

where U_t is unemployment and Δy_t is change in (log) GNP.

The identifying assumption is that u_t^d do not have permanent effects on the *level* of *GNP*

Computing long run effects

Compute long run (level) effects

$$\left[\begin{array}{c} U_t \\ \Delta y_t \end{array}\right] = \Phi_1 \left[\begin{array}{c} U_{t-1} \\ \Delta y_{t-1} \end{array}\right] + A_0^{-1}\mathbf{u}_t$$

by summing all future changes in GNP

$$\begin{aligned} & E\left[\sum_{s=0}^{\infty}\left[\begin{array}{c} U_{t+s} \\ \Delta y_{t+s} \end{array}\right] \mid \mathbf{u}_t\right] \\ = \; & \left[A_0^{-1} + \Phi_1 A_0^{-1} + \Phi_1^2 A_0^{-1} + \ldots \Phi_1^{\infty} A_0^{-1}\right]\mathbf{u}_t \\ = \; & \left(I - \Phi_1\right)^{-1} A_0^{-1}\mathbf{u}_t \end{aligned}$$

Linking the reduced form estimates to structural coefficients

As before, data can give us reduced form estimates of $\widehat{\Phi}_1$ and $\widehat{\Omega}$:

$$\begin{aligned}\widehat{\Phi}_1 &= \sum Y_t Y_{t-1}' \left[\sum Y_{t-1} Y_{t-1}'\right]^{-1} \\ \widehat{\Omega} &= \frac{1}{T-p} \sum \left(Y_t - \widehat{\Phi}_1 Y_{t-1}\right)\left(Y_t - \widehat{\Phi}_1 \mathbf{y}_{t-1}\right)'\end{aligned}$$

where $\Omega = E\left(\varepsilon_t \varepsilon_t'\right) = A_0^{-1}\left(A_0^{-1}\right)'$ and $Y_t \equiv [U_t \quad \Delta y_t]'$.

We now have everything we need to impose our identifying assumption that $\left(I - \Phi_1\right)^{-1} A_0^{-1}$ is upper triangular, i.e.

$$\left(I - \Phi_1\right)^{-1} A_0^{-1} = \begin{bmatrix} x & x \\ 0 & x \end{bmatrix}$$

Imposing the identifying assumption in practice

We want to find a matrix A_0^{-1} such that

$$\begin{aligned}(I-\Phi)^{-1} A_0^{-1} &= \begin{bmatrix} x & x \\ 0 & x \end{bmatrix} \\ A_0^{-1}\left(A_0^{-1}\right)' &= \Omega\end{aligned}$$

The best way to find A_0 is to again use the Choleski decomposition, but now of the matrix Q defined as

$$\begin{aligned}Q &\equiv (I-\Phi)^{-1} A_0^{-1}\left(A_0^{-1}\right)'\left((I-\Phi)^{-1}\right)' \\ &= (I-\Phi)^{-1}\,\Omega\left((I-\Phi)^{-1}\right)'\end{aligned}$$

so that

$$\text{chol}(Q) = (I-\Phi)^{-1} A_0^{-1}$$

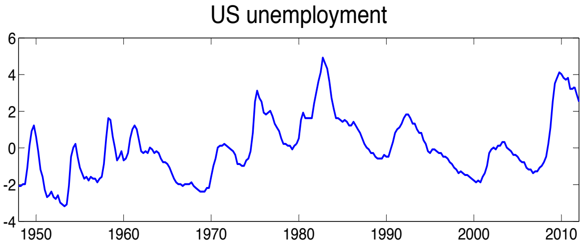
or that

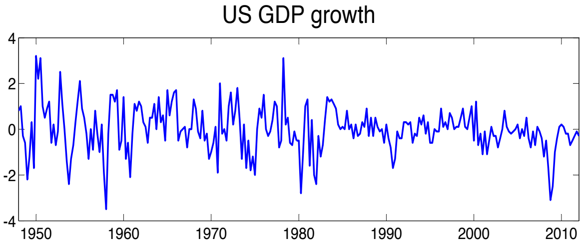
$$A_0^{-1} = (I-\Phi)\,\text{chol}(Q)$$

A procedure for using long run restrictions

1. Look at the data (Always useful....)
2. Estimate reduced form
3. Determine VAR order of reduced form
4. Find long run response matrix
5. Impose identifying assumption to find A_0^{-1}

Step 1: Look at the data





Step 2 and 3

Estimate reduced form

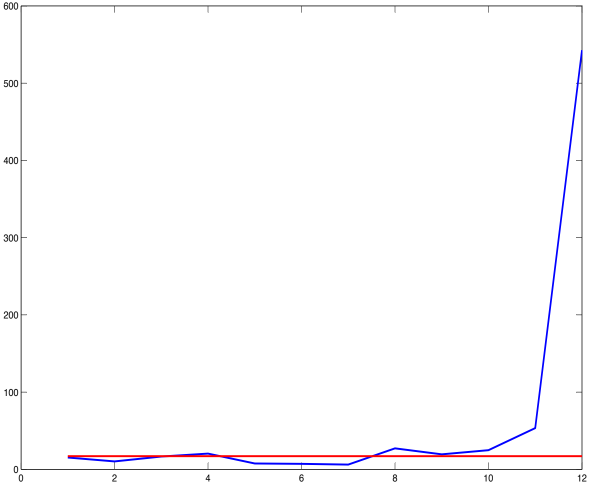
- Standard OLS

Determine VAR order of reduced form

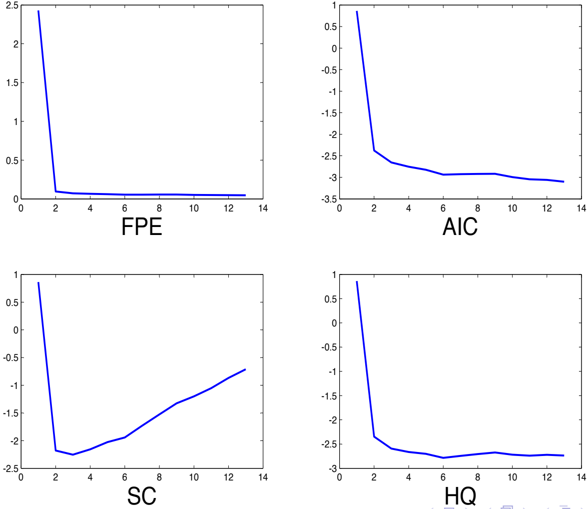
- Use LR test or FPE, AIC, HQ or Schwarz criterion

We know how to do all this, but let's look at plots.

Determine VAR order: LR tests



Determine VAR order



Step 4: Compute long run response

Computing the long run response in companion form

$$\xi_t = F\xi_{t-1} + Cu_t : \xi_t \equiv \begin{bmatrix} Y_t' & Y_{t-1}' & \cdots & Y_{t-p+1}' \end{bmatrix}'$$

and define companion matrix F as

$$F \equiv \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}, C = \begin{bmatrix} A_0^{-1} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

Long run response now given by

$$\begin{bmatrix} I_n & 0 \end{bmatrix} [I - F]^{-1} C$$

Step 4: Compute long run response *cont'd*

Long run response now given by

$$\left[\begin{array}{cc} I_K & 0 \end{array}\right] [I - F]^{-1} C$$

Define

$$\begin{aligned} W &\equiv \left[\begin{array}{cc} W_{11} & W_{12} \\ W_{21} & W_{22} \end{array}\right] \\ &= [I - F]^{-1} \end{aligned}$$

then long run response of $Y_t = W_{11} A_0^{-1}$ so that $A_0^{-1} = W_{11}^{-1} \text{chol}(W_{11} \widehat{\Omega} W_{11}')$.

Impulse response functions

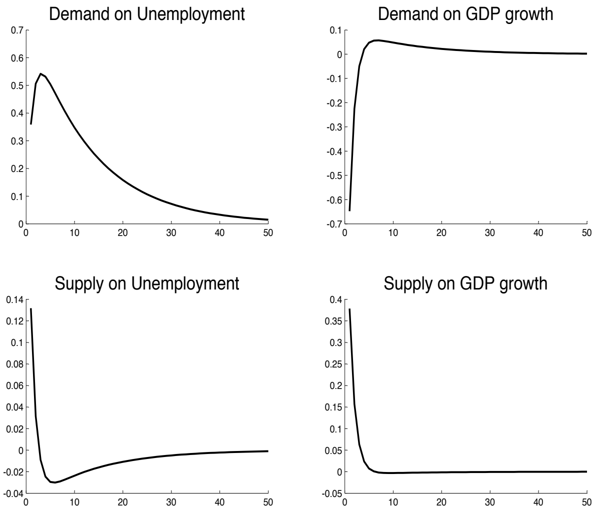
Put identified matrices in the appropriate places of the companion form in order to find impulse responses

$$\frac{\partial \xi_{t+s}}{\partial u_t} = A^s C$$

Since VAR order suggestions where somewhat ambiguous we should check for several different p.

- But for simplicity we will use $p = 1$

IRFs with p=1


Demand on Unemployment
Demand on GDP growth
Supply on Unemployment
Supply on GDP growth
0
10
20
30
40
50

A way to double check your IRFs:

What do we know about the impulse responses of first differences of GNP with respect to demand shocks?

- IRFs in growth rates should accumulate to zero, that is, area above and below zero line of impulse response of Δy_t with respect to a demand shock should "cancel".

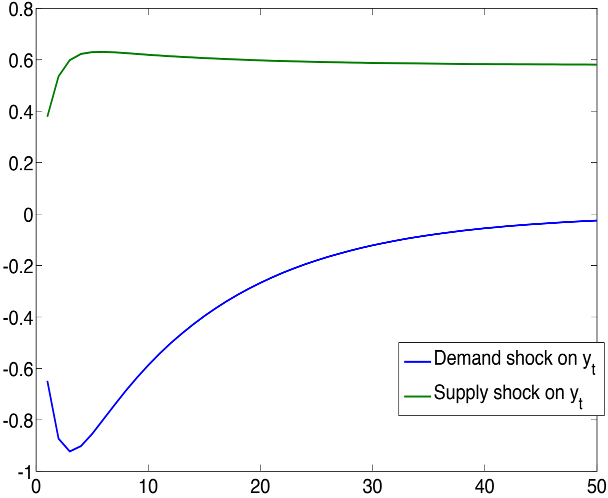
This can be checked by calculating the cumulative sums of the IRFs, i.e. for each s compute

$$\sum_{s=0}^{S} A^s C$$

then the second row, first column element should tend to zero as $S \to \infty$.

- It is easier to see in a graph

Cumulative IRFs with p=1



Variance decomposition

1. Start by computing the unconditional variance

$$\Sigma_{yy} = \Phi\Sigma_y\Phi' + CC' = \begin{bmatrix} \sigma^2_{\Delta y} & \sigma_{\Delta yu} \\ \sigma_{u\Delta y} & \sigma^2_u \end{bmatrix}$$

2. Compute variance of first variable using c_1c_1' instead of CC' in

$$\begin{aligned} \Sigma_{y1} &= \Phi\Sigma_{y1}\Phi' + c_1c_1' \\ &= \begin{bmatrix} \sigma^2_{\Delta y1} & \sigma_{\Delta yu1} \\ \sigma_{u\Delta y1} & \sigma^2_{u1} \end{bmatrix} \end{aligned}$$

3. Divide the resulting diagonal elements with corresponding diagonal element of unconditional covariance matrix, i.e. compute the fractions $\frac{\sigma^2_{\Delta y1}}{\sigma^2_{\Delta y}}$ and $\frac{\sigma^2_{u1}}{\sigma^2_{u}}$ to get the fraction of the unconditional variances explained by the first shock.
4. Repeat step 2-3 for each shock

Decomposing the variance in the Blanchard-Quah example

	u_t^d	u_t^s
U_t	0.99	0.01
Δy_t	0.74	0.26

- Most of the variance of both variables are explained by demand shocks
  - For unemployment, demand shocks explain virtually all of the variation

Check that decompositions sum to unity!

Historical Decompositions

What shocks at what time contributed to the business cycle during each moment in the sample?

$$
\begin{aligned}
Y_t &= \Phi Y_{t-1} + C\mathbf{u}_t \\
&= \Phi^t Y_0 + \Phi^{t-1} C\mathbf{u}_1 + \ldots + \Phi^{t-s} C\mathbf{u}_s + \ldots + \Phi^{t-t+1} C\mathbf{u}_{t-1} + C\mathbf{u}_t
\end{aligned}
$$

Decompose into the effect of each shock in period t as

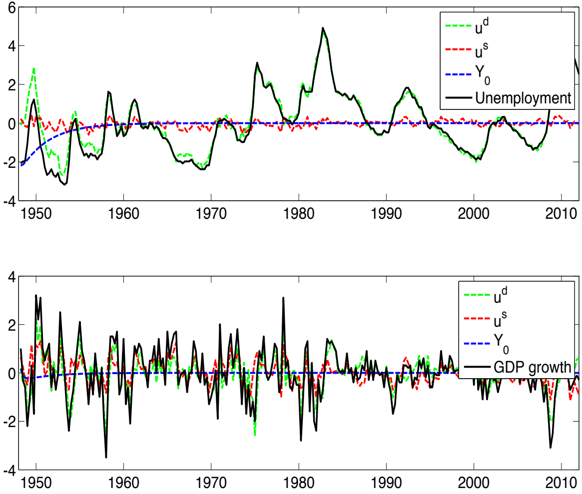
$$
\begin{aligned}
Y_t = \; & \Phi^t Y_0 + \Phi^{t-1}\mathbf{c}_1 u_{1,1} + \ldots + \Phi^{t-t+1}\mathbf{c}_1 \mathbf{u}_{1,t-1} + \mathbf{c}_1 u_{1,t} \\
& + \Phi^{t-1}\mathbf{c}_2 u_{2,1} + \ldots + \Phi^{t-t+1}\mathbf{c}_2 u_{2,t-1} + \mathbf{c}_2 u_{2,t}
\end{aligned}
$$

We can compute this for each t if we have a time series for the $\mathbf{u}_t$ given by

$$\mathbf{u}_t = C^{-1}\left(Y_t - \Phi Y_{t-1}\right)$$

It is illustrative to plot all this in one graph, containing $y_{1t}, y_{11,t}$ and $y_{12,t}$ and initial conditions effects.

Historical decomposition



Other examples of long run restrictions

Gali (1999) uses long run restrictions to analyze the implications of permanent productivity shocks for employment and hours worked.

- Identifying assumption: Only productivity shocks can have permanent effects on level of output
- Finds that identified productivity shocks do not cause an increase in hours worked or employment, as suggested by RBC models
- Gali argues that this is evidence in favor of New Keynesian models with sticky prices.
  - NK models predict that hours fall after productivity shock

More on this after the break....

A warning:

The Matlab command `R=chol(Q)` gives you an upper triangular matrix R such that $R'R = Q$ and **not** an R such that $RR' = Q$.

- I.e. R' is lower triangular.
- To get an upper triangular R' such that $R'R = Q$ use `R=chol(Q,'lower')`

- **Make sure to verify that your identified matrix A_0^{-1} satisfies $A_0^{-1}\left(A_0^{-1}\right)' = \Omega$!**

Two critiques of SVARs

1. Rudebusch vs Sims
2. Minnesota vs Long Run restrictions

Rudebusch vs Sims

Rudebusch (1998) argues that SVAR measures of monetary policy do not make sense

Sims argues that they do

Rudebusch vs Sims

What is the object we are discussing?

$$A_0 \begin{bmatrix} r_t \\ \Delta y_t \\ \pi_t \end{bmatrix} = A_1 \begin{bmatrix} r_{t-1} \\ \Delta y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \ldots + A_p \begin{bmatrix} r_{t-p} \\ \Delta y_{t-p} \\ \pi_{t-p} \end{bmatrix} + u_t$$

or in reduced form

$$\begin{bmatrix} r_t \\ \Delta y_t \\ \pi_t \end{bmatrix} = \Phi_1 \begin{bmatrix} r_{t-1} \\ \Delta y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \ldots + \Phi_p \begin{bmatrix} r_{t-p} \\ \Delta y_{t-p} \\ \pi_{t-p} \end{bmatrix} + \varepsilon_t$$

Sometimes there are many more than 3 variables included.

Rudebusch's argument:

- The interest rate equation does not look like the reaction function of a central bank
  - In reality, reaction functions are not linear and stable over time
  - There are too many lags for it to be a reasonable description of central bank decision making
  - Use of final (i.e. revised) data
- Monetary policy shocks look very different depending on which variables that are included in the regression, so how can we expect impulse responses to tell us anything about real monetary policy?

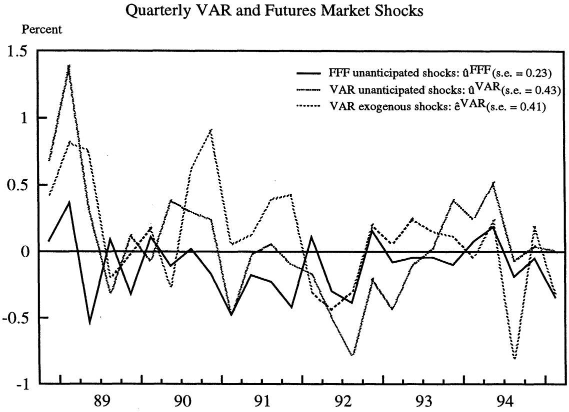

Quarterly VAR and Futures Market Shocks
Percent
1.5
1
0.5
0
-0.5
-1
FFF unanticipated shocks: $\hat{u}^{FFF}$(s.e. = 0.23)
VAR unanticipated shocks: $\hat{u}^{VAR}$(s.e. = 0.43)
VAR exogenous shocks: $\hat{e}^{VAR}$(s.e. = 0.41)
89
90
91
92
93
94

FIGURE 6

Sims' counter argument

- Linearity and time invariance are always approximations and this is a problem common to all macroeconomic models. Non-linearity and time varying rules add little explanatory power, though.
- Long lags in "reaction function" is just a statistical summary that does not imply that the Fed responds to "old information".
- Revised data: Can be handled by restricting the response of interest rates to only variables that are observed at the time of the decision.
- Sims has a subtle but important point about the how models can disagree about the shocks but agree about the effects of monetary policy.

Chari, Kehoe and McGrattan (JME2008)

Some background:

- SVARs using long run restrictions (e.g. Gali 1999) similar to Blanchard and Quah's (1989) find that hours worked decrease in response to permanent productivity shocks
- This is bad news for RBC models since they imply that hours should increase in response to productivity
- Conclusion: Only models fitting this pattern (i.e. sticky price models) are worth pursuing.

Chari et al challenges this conclusion

Chari, Kehoe and McGrattan (JME2008)

Chari et al shows that when the SVAR methodology is applied to data generated from a prototypical RBC model, hours appear to respond negatively to a productivity shock, even though in the model, they respond positively.

- This seems to suggest that the SVAR literature do not identify the productivity shock correctly

Chari et al argues that there are two sources that bias the results of the SVAR literature:

1. Small sample bias in VARs
2. Lag truncation bias

Chari, Kehoe and McGrattan (JME2008)

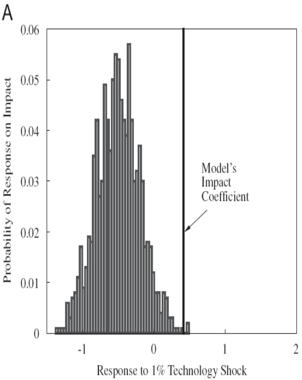
Lag truncation bias:
The structural model can be written as an $MA(\infty)$

$$\begin{aligned} Y_t &= A_0\varepsilon_t + A_1\varepsilon_{t-1} + ... + A_s\varepsilon_{t-s} : s \to \infty \\ &\equiv A(L)\varepsilon_t \end{aligned}$$

An implicit assumption in the SVAR literature is that $A(L)^{-1}$ exists and is equal to $I - \sum_{i=1}^{p} B_i L^i$ so that there exists a finite order VAR

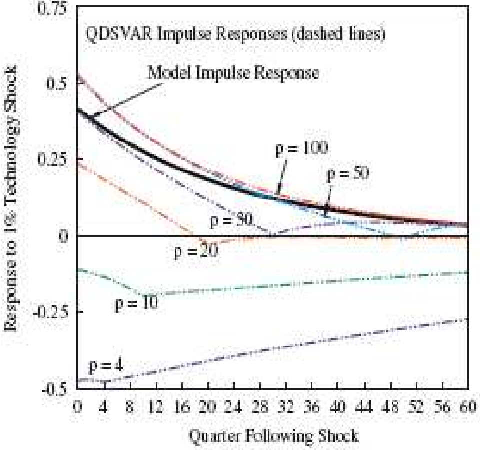
$$[I - \sum_{i=1}^{p} B_i L^i] Y_t = \varepsilon_t$$

where p is a low number (typically $p = 4$).



A
0.06
0.05
0.04
0.03
0.02
0.01
0
Probability of Response on Impact
Model's Impact Coefficient
-1
0
1
2
Response to 1% Technology Shock
B
0.06
0.05
0.04
0.03
0.02
0.01
0
Probability of Response on Impact
Model's Impact Coefficient
-1
0
1
2
Response to 1% Technology Shock

A



0.75
0.5
0.25
0
-0.25
-0.5
QDSVAR Impulse Responses (dashed lines)
Model Impulse Response
p = 100
p = 50
p = 30
p = 20
p = 10
p = 4
Response to 1% Technology Shock
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60
Quarter Following Shock

Suggested strategy to use SVARs to guide model design

- It is risky to compare estimated IRFs to true model IRFs
- A better strategy is to compare estimated IRFs from actual data with estimated IRFs from data generated by theoretical model

The Minnesota slogan is *Do to the model what you do to the data*
That is reasonable advice

- Also applies to de-trending issues etc