Assessing Structural VAR’s

by

Lawrence J. Christiano, Martin Eichenbaum and Robert Vigfusson

May 2005
Traditions for constructing GE models.

• First tradition focusses on at most a handful of key shocks, deliberately abstracts from smaller shocks.
  – Classic example: Kydland and Prescott.
  – Conundrum: how do you empirically evaluate models (which contain only a subset of the shocks) with the data (which are driven by all the shocks)
  – Structural VARs have potential to provide resolution to conundrum
    * Assess empirical performance of model relative to a particular set of shocks.

• Second tradition
  – Build macro models with large numbers of shocks.- complete characterization of DGP.
  – Avoids KP conundrum.
  – SVAR’s still useful for diagnostic purposes.
Questions That We Investigate

- What are bias properties of VAR based impulse response estimators?
- What are bias properties of standard estimators of sampling uncertainty in the estimator?
- Are there easy to implement variants of standard procedures which improve bias properties of response function estimators?

- We address these questions using data generated from dynamic GE models.
  - Look at Long Run Restrictions and Short Run Restrictions

- Our conclusion:
  - Structural VARs provide valuable information for building empirically plausible models of aggregate fluctuations.
Challenges

• Growing literature calls into question ability of VARs to uncover dynamic response of macroeconomic variables to structural shocks.
  – Focuses on long run based identification schemes

• Intellectual father of this literature - Sims (1972)
  – Problem: need a reliable estimate of sum of coefficients in distributed lag regressions.
  – Hard to get this even if individual coefficients are reasonably precisely estimated.
  – See also Faust and Leeper and Pagan.

• More recently EGG and CKM examine reliability of VAR-based inference using long run identifying restrictions.
  – CKM are exceedingly critical.
Findings for Short Run Restrictions

• DGP: variants of a standard real business cycle model augmented by timing restrictions.
  – Focus on response of hours to technology shock.

• Conclusion:
  – VAR’s perform remarkably well
  – Virtually no bias - either in point estimates or estimates of sampling uncertainty.

• Very comforting for vast literature that uses short run restrictions to identify consequences of shocks to economy.
Findings for long run restrictions

• When technology shocks account for a substantial fraction of business cycle fluctuations in output, VAR based analysis is reliable.
  – Some evidence of bias when tech shocks play much smaller role relative to estimates in standard RBC literature.

• First way to eliminate bias:
  – When number of variables in VAR exceeds number of important driving shocks, bias in impulse response estimators is substantially reduced.
  – Widespread consensus: only a handful of important shocks drive aggregate fluctuations

  – A reasonably small number of shocks will reduce small sample bias even when technology shocks are relatively unimportant.
Second way to eliminate bias:

- Integrate a non-parametric estimator of zero-frequency spectral density of data.

- Identification based on long run restrictions involves estimating features of spectral density of the data at frequency zero.
  - Estimator implicit in standard VAR-based methods - poor properties when technology shocks are relatively unimportant in output fluctuations.

- We propose / implement an adjustment to standard, pure VAR-based methods
  - Newey-West non-parametric estimator of zero-frequency spectral density

- With this adjustment, there’s no substantial bias in estimated impulse response functions, for any of the cases that we consider.
Relationship to CKM

• Most of CKM’s analysis:
  – Consequences of working with the first difference of hours worked, when the levels specification is the correct one. This is uninteresting.

• CKM argue that when hours worked enter analysis in levels, there’s enormous sampling uncertainty in estimated impulse response function.
  – So VARs are essentially useless.

• We reach a different conclusion, for three reasons

• Reason 1
  – CKM don’t consider short run restrictions: only consider case of long-run restrictions.
Relationship to CKM ...

• Reason 2:

  – CKM reach their conclusions based on two ‘worst case’ examples.
    * Technology shocks account for only a very small proportion (23 and 17 percent!!, respectively) of output fluctuations
    * Number of variables included in VAR is equal to the number of important shocks in model.

• We also find evidence of bias in experiments sharing these properties.

• Bias is easily dealt with, doesn’t reflect a fundamental flaw in structural VARs.
Relationship to CKM ...

• CKM conjecture:
  – Distortions in their two examples are due to omission of capital in the VAR.

• Conjecture is false
  – We don’t include capital in our VAR
  – Even so there’s virtually no bias when we use our adjusted VAR estimation procedure in their two examples.

  – Even with standard VAR procedures there’s no bias when technology shocks play a more important role in aggregate fluctuations than they do in CKM’s examples.
Reason 3 for disagreeing with CKM

- CKM strategy for implementing long-run restrictions is ‘anomalous’ relative to literature.
  - CKM: A positive technology shock is one that drives labor productivity up in the period of the shock.

- Standard approach
  - A positive technology shock is one that drives labor productivity up in the long run.
Reason 3 for disagreeing with CKM ...

- CKM’s procedure leads to a bimodal small sample distribution in response of hours to a technology shock.
  
  - About 20 percent of the time, in their model hours worked initially falls, while labor productivity falls permanently despite rising initially.
  - They identify these shocks as a *positive* technology shock.
  
  - Standard strategy would identify these responses as consequences of a negative technology shock.

- CKM’s ‘anomalous’ identification strategy leads them to overstate sampling uncertainty by a factor of two.
Outline

• Analysis Based on Simple RBC Model

  – Long Run Restrictions
  – Short Run Restrictions

• Reconciling with CKM

• Concluding Comments
DGP: A Generic RBC Model

- Preferences:

\[
E_0 \sum_{t=0}^{\infty} (\beta (1 + \gamma))^t \left[ \log c_t + \psi \frac{(\bar{l} - l_t)^{1-\sigma}}{1-\sigma} \right].
\]

- Constraints:

\[
c_t + (1 + \tau_{x,t}) [(1 + \gamma) k_{t+1} - (1 - \delta) k_t] \leq (1 - \tau_{lt}) \omega_t l_t + r_t k_t + T_t.
\]

\[
c_t + (1 + \gamma) k_{t+1} - (1 - \delta) k_t \leq k_t^{\alpha} (Z_t l_t)^{1-\alpha}.
\]

- Shocks:

\[
\log z_t = \mu_Z + \sigma_z \varepsilon_t^z
\]
\[
\tau_{lt+1} = (1 - \rho_l) \bar{\tau}_l + \rho_l \tau_{lt} + \sigma_l \varepsilon_t^{l},
\]
\[
\tau_{xt+1} = (1 - \rho_x) \bar{\tau}_x + \rho_x \tau_{xt} + \sigma_x \varepsilon_t^{x}
\]
Two Versions of Model

- Differentiated by timing assumptions.

- **Standard version**
  - All time $t$ decisions taken after realization of the time $t$ shocks.

- **Recursive version**
  - First, $\tau_{lt}$ is observed, after which labor decisions are made.
  - Second, other shocks are realized.
  - Then, agents make their investment and consumption decisions.
  - Finally, labor, investment, consumption, and output occur.
Standard Version of Model

• $\varepsilon_t^z$ is only shock that has a permanent impact on output and labor productivity

\[ a_t \equiv \frac{y_t}{l_t}. \]

• Exclusion restriction:

\[ \lim_{j \to \infty} [E_t a_{t+j} - E_{t-1} a_{t+j}] = f(\varepsilon_t^z \text{ only}), \]

• Sign restriction
  – $f$ is an increasing function.
Estimating Effects of a Positive Technology Shock

- Impose exclusion and sign restrictions on VAR to compute $\varepsilon_t^z$, identify its dynamic effects on macroeconomic variables.

\[
Y_{t+1} = B(L)Y_t + u_{t+1},\ Eu_tu'_t = V,
\]

\[
B(L) \equiv B_1 + B_2L + \ldots + B_pL^{p-1},
\]

\[
u_t = C\varepsilon_t,\ E\varepsilon_t\varepsilon'_t = I,\ CC' = V
\]

\[
Y_t = \begin{pmatrix}
\Delta \log a_t \\
\log l_t \\
x_t
\end{pmatrix},\ C = [C_1:C_2:C_3],\ e_t = \begin{pmatrix}
e_{1t} \\
e_{2t} \\
e_{3t}
\end{pmatrix},\ a_t = \frac{Y_t}{l_t}
\]
Estimating Effects of a Positive Technology Shock ...

- Impulse Response to Positive Technology Shock \((e_{1t})\):

\[ \tau [I - B(1)]^{-1} C \varepsilon_t, \]

- \(\tau\): a row vector with all zeros, except unity in the first location.
- \(B(1)\): sum, \(B_1 + \ldots + B_p\).
- \(\tilde{E}_t\): expectation operator, conditional on \(\tilde{\Omega}_t = \{Y_t, \ldots, Y_{t-p+1}\}\).

- To compute the dynamic effects of \(\varepsilon^z_t\), we require \(B_1, \ldots, B_p\) and \(C_1\), the first column of \(C\).
Identification Problem

- From applying OLS to both equations in VAR we ‘know’:

\[ B_1, ..., B_p, V \]

- Problem, Need first Column of \( C, C_1 \)

\[ CC' = V \]

- Identification Problem:

Not Enough Restrictions to Pin Down \( C_1 \)

- Need More Restrictions
Identification Problem ...

- Two Key Properties of DGP:

  - Long-Run Exclusion Restriction:

    \[ \lim_{j \to \infty} [E_t a_{t+j} - E_{t-1} a_{t+j}] = f(\varepsilon^z_t \text{ only}) \]

  - Sign Restriction:

    \[ f \text{ increasing in } \varepsilon^z_t \]

- These Properties Provide Sufficient Restrictions to Pin Down $C_1$ (uniquely)
Standard Algorithm for Computing $C_1$

- Step 1: Compute Lower Triangular Choleski Decomposition, $D$

\[
DD' = [I - B(1)]^{-1} V [I - B(1)']^{-1}
\]

subject to $D(1, 1) > 0$.

- Step 2: Solve

\[
C = [I - B(1)] D.
\]

- Remark: this $C$ Satisfies all Restrictions

\[
CC' = [I - B(1)] DD' [I - B(1)'] = V
\]

(exclusion restriction) \( [I - B(1)]^{-1} C = \begin{bmatrix} x & 0, \ldots, 0 \\ \text{numbers} & \text{numbers} \end{bmatrix} \)

(sign restriction) $x > 0$
Standard Algorithm for Computing $C_1$ ...

- Recall:

$$DD' = [I - B(1)]^{-1} V [I - B(1)']^{-1}, \quad C = [I - B(1)] D.$$

- Remark:

  - Sign of $C(1, 1)$ unrestricted
  - Positive Technology Shock is One that Leads to a Permanent Rise in Productivity
  - Could Lead to Contemporaneous Drop in Productivity
The Recursive Version of the Model

- Long-run identification strategy outlined above can be rationalized in this model.

- Alternative procedure for identifying $\varepsilon_t$ that does not rely on estimating long-run responses to shocks can also be rationalized.

- ‘Short run’ strategy: involves recovering $\varepsilon_t$ using just realized one-step-ahead forecast errors in labor productivity and hours, as well as the second moment properties of those forecast errors.
The Recursive Version of the Model ...

\[ Y_t = \begin{pmatrix} \log l_t \\ \Delta \log a_t \\ x_t \end{pmatrix}, \]

\[ u_t = \begin{pmatrix} u_t^l \\ u_t^a \\ u_t^x \end{pmatrix} \]

\[ u_t = C\varepsilon_t, \quad E\varepsilon_t\varepsilon_t' = I, \quad CC'' = V \]

- To compute dynamic response of \( Y_t \) to \( \varepsilon_t \), need \( B_1, \ldots, B_q \) and second column of \( C \).
  - Compute \( CC'' = V \), where \( C \) is lower triangular Choleski decomposition of \( V \).
    Take second column of that matrix.
Parameterizing the Model

- Different versions of the RBC model, distinguished by the nature of exogenous shocks.
- As in CKM we assume 

\[
\begin{align*}
\beta &= 0.9722^{1/4}, \quad \theta = 0.35, \quad \delta = 1 - (1 - 0.0464)^{1/4}, \\
\psi &= 2.24, \quad \gamma = 1.015^{1/4} - 1, \quad \bar{l} = 1300, \\
\bar{\tau}_x &= 0.3, \quad \bar{\tau}_l = 0.27388, \quad \mu_z = 1.016^{1/4} - 1, \quad \sigma = 1.
\end{align*}
\]

**KP Specification**

- Technology shock process: Prescott (1986):

\[
\log z_t = \mu_Z + 0.011738 \times \varepsilon_t^z.
\]

- EGG (2005) update Prescott’s analysis, estimate \( \sigma_z \) to be 0.0148.
Parameterizing the Model ...

– To be conservative, we use Prescott’s estimate.

• Law of motion for $\tau_{l,t}$ as follows.
  – Household / firm FONC’s imply:

\[
\tau_{l,t} = 1 - \frac{c_t}{y_t} \frac{l_t}{l_t 1 - \theta} \psi.
\]

\[
\tau_{l,t} = (1 - 0.9934) \times 0.2660 + 0.9934 \times \tau_{l,t-1} + 0.0062 \times \varepsilon_l^t.
\]

• Percent of variance in HP-filtered, log output due to technology shocks is 73%.
  – Consistent with key claim of KP.
Parameterizing the Model ...

CKM Benchmark Specification

\[ \log z_t = \mu_Z + \log z_t = \mu_Z + 0.00581 \times \varepsilon_t^z \]
\[ \tau_{lt} = (1 - \rho_l) \bar{\tau}_l + \rho_l \tau_{l,t-1} + 0.00764 \times \varepsilon_{t}^l, \quad \rho_l = 0.93782. \]

• Percent of variance in HP-filtered, log output due to technology shocks is only 23%.
  – CKM use this specification to criticize Gali (1999).
  – Embodies Gali’s main hypothesis that technology shocks play only a very small role in business cycle fluctuations.

Other Specifications

• Vary \( \sigma \) and \( \sigma_l \)
  – Important quantitative effect on contribution of technology shocks to volatility of output.
Adding Shocks and Variables to model.

Three Shocks, Two Important Specification

\[ \tau_{xt} = \bar{\tau}_x + 0.0001 \times \varepsilon_t^x \]

- Three-variable VAR, with \( x_t = \log \left( \frac{c_t}{y_t} \right) \)

Three Shocks, Three Important Specification

- As in CKM

\[ \tau_{xt} = (1 - 0.9) \bar{\tau}_x + 0.9 \times \tau_{x,t-1} + 0.01\varepsilon_t^x. \]

\[ x_t = \begin{pmatrix} \log \left( \frac{c_t}{y_t} \right) \\ \tau_{xt} + w_t \end{pmatrix}, \]

\[ w_t \sim \mathcal{N}(0, 0.0001) \]
Experiments

- Simulate 1000 data sets, each of length 180 observations, using GE model as DGP.
  - Shocks $\varepsilon_t^\zeta$, $\varepsilon_t^l$ and possibly $\varepsilon_t^x$ are drawn from $i.i.d.$ standard normal distributions.

- Estimate a four lag VAR.
  - Report Mean Impulse Response Function over 1000 synthetic data sets.
  - Measure of sampling uncertainty associated with the estimated dynamic response functions.
    * Calculate standard deviation of points in estimated impulse response functions across the 1000 synthetic data sets (Grey Area).
    * Also calculate top 2.5% and bottom 2.5% of the estimated coefficients in dynamic response functions across the 1000 synthetic data sets (Red lines).
Experiments ...

- Would an econometrician correctly estimate true uncertainty associated with estimated dynamic response functions.

  - For a given synthetic data set, estimate a VAR and use it as data generating process to construct 200 synthetic data sets, each of length 180 observations?
  - For each synthetic data set, estimate a new VAR and impulse response function.
  - Calculate standard deviation of coefficients in the impulse response functions across the 200 data sets.
  - Take average of these standard deviation across the 1000 synthetic data sets that were generated using the economic model as the data generating process (Line with 0’s).
Figure 2: Analysis of Short–Run Identification Assumption

KP Model

Benchmark

CKM Model

Note: hours response, in percent terms, to a 1.2 (KP) or 0.6 (CKM) percent innovation in technology, $Z_t$.
Solid line – mean response, Gray area – mean response plus/minus two standard errors,
Starred line – true response, Dashed line – 95.5 percent probability interval of responses,
Circles – average value of econometrician estimated plus/minus two standard errors.
Summary of Findings with Short Run Restrictions

- No evidence whatsoever of bias in the estimated impulse response functions.

- An econometrician wouldn’t be misled in inference using standard procedures for constructing confidence intervals.

- SVAR’s perform remarkably well.
  – If relevant recursive assumptions are satisfied in DGP, standard SVAR procedures will reliably uncover and identify the dynamic effects of shocks to the economy.

- We did not include capital as a variable in the VAR.
  – Claims in CKM to contrary, omitting economically relevant state variable capital does not in and of itself pose a problem for inference using SVAR’s.
Figure 3: Analysis with Kydland–Prescott Specification

- **Standard Estimator**
- **Benchmark KP**
- **Newey–West Spectral Estimator**

**Note:** hours response, in percent terms, to a 1.2 percent innovation in technology, $Z_t$.

- Solid line – mean response.
- Gray area – mean response plus/minus two standard errors.
- Starred line – true response.
- Dashed line – 95.5 percent probability interval of responses.
- Circles – average value of econometrician estimated plus/minus two standard errors.
Long Run Restrictions: KP Specification

- Virtually no bias in point estimates.
- Considerable sampling uncertainty, but econometrician wouldn’t be misled with respect to inference.

- Hansen Indivisible Labor model, $\sigma = 0.0001$.
  - Bias associated with estimator increases (very) slightly.
    * Percent of variance in HP-filtered, log output due to technology shocks is 62%.
  - Econometrician wouldn’t be misled about sampling uncertainty.

- EGG: $\sigma = 1.24$.
  - Bias almost disappears, and the sampling uncertainty shrinks drastically.
  - Percent of variance in HP-filtered, log output due to technology shocks is 92%.
Long Run Restrictions: KP Specification ...

- Three variable, three shock version of model.
  - Noticeable degree of bias associated with the estimated impulse response function.
    * But relatively small in relation to the sampling alternative.
    * Econometrician’s estimated confidence interval is roughly correct, on average.
    * Percent of variance in HP-filtered, log output due to technology shocks is 57%
Why Does Bias Appear in Last Case?

- Sims (1972) : can characterize the VAR parameter estimates econometrician would obtain in large sample ($\hat{B}_1, ..., \hat{B}_q$ and $\hat{V}$)

\[
\hat{V} = V + \min_{\hat{B}_1, ..., \hat{B}_q} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[ B(e^{-i\omega}) - \hat{B}(e^{-i\omega}) \right] S_Y(\omega) \left[ B(e^{i\omega}) - \hat{B}(e^{i\omega}) \right]' d\omega
\]

- $S_Y(\omega)$ is associated spectral density, at frequency $\omega$.

- Econometrician chooses VAR lag matrices to minimize a quadratic form in difference between estimated and true lag matrices
  - Assigns greatest weight to frequencies where spectral density is greatest.
  - If there’s specification error, then $\hat{B} \neq B$ and $V > \hat{V}$.

- There’s specification error because true VAR implied by models has $q = \infty$, but econometrician uses a finite value of $q$. 
Why Does Bias Appear in Last Case? ...

- Two key ingredients to computing impact effects of shocks:
  - Estimate of variance covariance matrix of VAR disturbances and spectral density of $Y_t$ at frequency zero.
  - Variance covariance matrix is likely to be estimated precisely.
  - Problem with spectral density at frequency zero.
    * Need sum of estimated VAR matrices.
    * No particular reason for this to be estimated precisely by ordinary least squares.
    * Sum of lag VAR matrices corresponds to $\omega = 0$ and least squares will pay attention to this only if $S_Y(\omega)$ happens to be relatively large in a neighborhood of $\omega = 0$.

- Replace $S_0$ with Newey-West estimator.

- Figure 3
  - Bias is reduced
  - Less sampling uncertainty.
Figure 3: Analysis with Kydland–Prescott Specification

Standard Estimator

Benchmark KP

Newey–West Spectral Estimator

Note: hours response, in percent terms, to a 1.2 percent innovation in technology, $Z_t$.

Solid line – mean response, Gray area – mean response plus/minus two standard errors,
Starred line – true response, Dashed line – 95.5 percent probability interval of responses,
Circles – average value of econometrician estimated plus/minus two standard errors.
CKM Long Run Results

- Benchmark CKM: substantial bias
- Disappears if you
  - Raise importance of technology shocks in output volatility (23% versus 54% versus 73%)
  - Adopt Newey-West Estimator of Spectrum at frequency zero.
  - Insert Figures 4 and 5
- Also bias is substantially reduced if you could one more variable than important shock
  - INSERT FIGURE 5.
Figure 4: Analysis with CKM Specification

Standard Estimator  
Newey–West Spectral Estimator

Benchmark CKM

CKM with Half the Volatility in the Labor Tax Shock

CKM with One-third the Volatility in the Labor Tax Shock

Note: hours response, in percent terms, to a 0.6 percent innovation in technology, $Z_t$.
Solid line – mean response, Gray area – mean response plus/minus two standard errors,
Starred line – true response, Dashed line – 95.5 percent probability interval of responses,
Circles – average value of econometrician estimated plus/minus two standard errors.
Note: hours response, in percent terms, to a 0.6 percent innovation in technology, $Z_t$.
Solid line – mean response, Gray area – mean response plus/minus two standard errors,
Starred line – true response, Dashed line – 95.5 percent probability interval of responses,
Circles – average value of econometrician estimated plus/minus two standard errors.
Figure 6: Analysis with CKM Specification: The Impact of Adding Variables and Shocks to the Benchmark

Note: hours response, in percent terms, to a 0.6 percent innovation in technology, $Z_t$.

Estimation results for Standard VAR estimator.

Solid line – mean response, Gray area – mean response plus/minus two standard errors,
Starred line – true response, Dashed line – 95.5 percent probability interval of responses,
Circles – average value of econometrician estimated plus/minus two standard errors.
CKM Long Run Results ...

- Long-run restrictions may work tolerably well, even in a world that poses as sharp a challenge for these methods as benchmark CKM model.
  
  – Encouraging because there’s widespread perception there is at most 2-4 important shocks driving the economy.

  – With 5 or more variables in the VAR, kind of problems inherent in the CKM benchmark model may not be severe even if you don’t use Newey West estimator and technology shocks aren’t very important.
CKM and Sampling Uncertainty

- CKM’s sign restriction leads them to confuse positive and negative technology shocks
  - Greatly overstate sampling uncertainty

- INSERT FIGURE 7
  - Bottom panel: impulses from simulations when we impose sign restriction.
    * In all cases productivity eventually rises,
    * But in some cases it \textit{initially} falls (20% of the time)
    * In these cases, initial hours response is especially strong.
  - Top panel: CKM restrictions.
    * If initial productivity falls - CKM call it a negative technology shock.
    * This is true even the shock leads to a permanent long run decline in productivity.
  - Responses of hours worked, which were strong positive numbers are interpreted to be strong negative numbers.
Figure 7: Identification Strategies

A Positive Technology Shock Increases Labor Productivity on Impact

Note: Responses, in percent terms, to a 0.6 percent innovation in technology, $Z_t$.

Estimation results for Standard VAR estimator using the Benchmark CKM model.
CKM and Sampling Uncertainty ...

- Result: CKM get a bimodal distribution for impact effects, falsely conclude variance is enormous.
  - On this basis they erroneously conclude VAR’s are uninformative.
Conclusion

- Studied properties of structural VAR’s for uncovering impulse response functions to shocks.
- With short run restrictions, VAR’s perform remarkably well.
- With long run restrictions, with one exception VAR’s also perform well.
  - Potentially perform less well when shock under investigation isn’t very important.
  - When there are enough variables in VAR, problems are greatly mitigated.
- Develop and implement a modified VAR approach
  - Leads to drastic improvement even when technology shocks play a limited role in aggregate fluctuations and a small number of variables are included in VAR.