590 Macroeconomics: Model Evaluation and VARs

Paul Beaudry and Franck Portier

March 2016

Estimation and evaluation of stochastic dynamic models

- **Outline of section**
- Start with a (brief) general overview of model estimation and evaluation
- Introduce vector autoregressions (VARs), and especially focus on different approach to the identification of structural impulse responses.
    - Can be used for model evalutaion or as target for estimation
- Discuss Bayesian estimation using Dynare (Colin)

General Framework for linear models

- The relations determining equilibrium outcomes for a macro-model can generally be written in the form

$$X_t = A_1 X_{t-1} + A_2 E_t[X_{t+1}] + A_3 Z_t$$

$$Z_t = \Gamma Z_{t-1} + \epsilon_t$$

where X_t is the set of endogenous variables, and Z_t is a set of exogenous variables, and ϵ_t is a vector of iid shocks.. This is sometimes called the structral form of the model. Since expectations are not usually directly obversable, such a system only becomes operational when augmented with a theory of expectation; the baseline case in modern macro being rational (or model consistent) expectation. However, other theories of expectation formation can be rather easily handled in same structure. Estimation and evaluation can be done directly on the above equation, but most often the equilibrium if first solved to obtain the reduced form or state space representation.

Reduced Form and State Space form

- Solving the model under RE, we get the form

$$X_t = AX_{t-1} + BZ_t$$

$$Z_t = \Gamma Z_{t-1} + \epsilon_t$$

- where the matrix A and B are functions of A_1, A_2, A_3 and most importantly under RE B is also a function Γ. The last observation being the essence of the Lucas critque. This is also refered to a the reduced form, with the recognition that the coeficients in B are not structural as they depend on Γ. For bring this to data, often system is written in pure state space form

$$\tilde{X}_t = \tilde{A}\tilde{X}_{t-1} + \tilde{B}\epsilon_t$$

$$Y_t = \tilde{X}_t + \nu_t$$

- where $\tilde{X}$ is simply the vector of X and Z stacked, $\tilde{A}$ depends on all the structural parameters including Γ, Y_t is a set of observable variables, and ν_t are measurement errors.

- For here, we will disregard the measurement problem and focus on either model in VAR form

$$\tilde{X}_t = \tilde{A}\tilde{X}_{t-1} + \tilde{B}\epsilon_t$$

- or in reduced form form

$$X_t = AX_{t-1} + B(A, \Gamma)Z_t$$

$$Z_t = \Gamma Z_{t-1} + \epsilon_t$$

- where the fact that $\tilde{A}$, A and B depend on a set of underlying parameters imposed cross-equation restrictions.
- Note: Easy to get an estimate of $\tilde{A}$, the difficulty is getting form these estimates to the estimates of the structural parameters. This is the identification problem.

What is a theory?

- A theory can be thought as simply a set of restrictions of the VAR representation of the data
- Since the VAR representation can also be expressed in terms of variances and co-varainces, a theory is simply a restriction on the co-varaince properties of the data!
- Two questions
  1. How to estimate the structural parameters of the model from observations
  2. How to evalute the 'appropriateness' of model
- First is direct econometrics
- Second can also be seen as direct econometrics, but here more economic judgement often applied

Estimation and evaluation

- Two main econemtric approaches
- Moment matching and FIML
- Moment matching includes GMM, SMM (see Chapter 7 in Dejong and Dave, Chapter 6 on calibration)
- Attractive aspects: allows one to focus on moments targeted by theory, but choice of moments may appear arbitrary (especially when looking at over-id tests.)
- Much of calibration can be thought as a moment matching exercise. First a just identified system, then an informal test of over-identification

Matching moments

- The GMM idea. If theory implies

$$E(f(z_t, \gamma)) = 0$$

- create

$$g(Z, \gamma) = \frac{1}{T} \sum_{T} f(z_t, \gamma)$$

- get estimate of γ so that $g(Z, \gamma)$ is as close as possible to zero, ie,

$$\min_{\gamma} g(Z, \gamma)' \Omega g(Z, \gamma)$$

- Ω is a positive definite matrix. There are preferred choices for Ω, thats is, produce efficient estimators.

Matching moments

- The GMM was first motivated by Euler equations (estimate on structural form)
- Now application has been greatly extended to allow one to choose parameters to match any implication of model. Idea: Take any difference between model and data which should be zero in the limit, and choose parameters to get that as close as possible.
- Example: Comparing co-variance implied by model versus data, even if the model implication needs to be simulated. This is simulated method of moments.
- This framework allows for testing when over-identified. Note: cross-equation restrictions playing a huge role in identification and testing. For identification, need the number of parameters to be less than the number of moments, and this needs to be strict for testing.
- The most "difficult" issue is getting the standard errors right on the structral parameters

Evaluating stochastic dynamic models

- Approach most often favored by econometricien: full system estimation: ex: maximum likelihood estimation and testing of all cross-equation restriction (over-identification restrictions)
- Drawback: model is a simplification of reality. Strong test will almost always be rejected, except if one favors highly parametrized model.
- In such cases, if we ignore the test, maximum likelihood is used to quantify the relative importance of different shocks. By estimation, we get estimates of the model parameters and variances of innovations which best fit the data. If just identified, or close to just identified, then we don't learn if model is good: it is by assumption.
- : Remember: identification refers to the possibility of extracting the parameters of the model from the observed co-variance structure.

Structural moving average representation approach: motivation

- The above discussion discussed how macro models can be estimated and evaluated
- However, such approaches are often not very informative about model fit.
  1. If they fit badly (ie overid rejected), often don't know why. Does the model work well on some dimension or not others. Maybe doesn't fit well because we forgot to inculde a shock.
  2. Even if model fits well, could be for the wrong reasons. By looking at data with all shocks considered together, may confounding effects (especially with limited moment approach)
- This motives looking at more economically informative methods. SVAR

Basic issues on VARs as a description tools

- A Var: vector auto-regression. From data, we can easily estimate from stationary data a relation of the form.

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_N Y_{t-N} + \epsilon_t$$

where Y_t and ϵ_t are vectors and ϵ_t is iid (enough lags), with variance co-variance matrix Σ.

- Alternatively, this can be written as

$$A(L) Y_t = \epsilon_t$$

where $A(L) = I + A_1 L + ... + A_N L^N$, is a polynomial in lag operator.

- Vars are a simple framework for making forecasts

Basic issues on VAR

- Vars can be inverted to get a moving average representation (either by simulation or algebra), ie,

$$Y_t = C(L)\epsilon_t$$

where $C(L) = I + C_1 L + C_2 L^2 + ...$ (Wold representation, or reduced form moving average)

- And can get impulse responses, example, for the the effect if ϵ_{1t} on Y_{1t},

$$\{1, c_{11}^1, c_{11}^2, c_{11}^3 ...\}$$

- But this does not give much information as the Σ is not diagonal and the shocks are meaningless (linear combinations of structural shocks).

Basic issues on VAR

- The innovations in a Var can always be thought as coming for a set orthogonal-normal shocks υ_t:

$$\epsilon_t = K\upsilon_t$$

- If fact there always exist a triangular matrix K such that $KK' = \Sigma$ (Choleski decomposition)
- Therefore we can always construct a orthogonal moving average

$$Y_t = \tilde{C}(L)\upsilon_t = C(L)K\upsilon_t$$

- With associated impulse responses that contain all the information, and we can do variance decompositions

Variance decomposition

- Consider bi-variate case

$$Y_{1t} = \tilde{c}^0_{11}\upsilon_{1t}+\tilde{c}^0_{12}\upsilon_{2t}+\tilde{c}^1_{11}\upsilon_{2t-1}+\tilde{c}^1_{12}\upsilon_{2t-1}+...++\tilde{c}^i_{11}\upsilon_{2t-i}+\tilde{c}^i_{12}\upsilon_{2t-i}+...$$

- The contribution of υ_{1t} to the variance of Y_{1t} at horizon N is given by

$$\frac{\sum_{i=0}^{N}(\tilde{c}^i_{11})^2}{\sum_{i=0}^{N}(\tilde{c}^i_{11})^2+\sum_{i=0}^{N}(\tilde{c}^i_{12})^2}$$

- Question: But what are these υs, do they relate to anything we care about.
- In general: no. But sometimes, maybe yes. This gets us to the structural var literature. Example with monetary shocks

Main identification issue

- From the data we have a Wold representation

$$Y_t = C(L)\epsilon_t$$

where $C(L) = I + C_1 L + ...$, ϵ_t iiid, and variance covariance matrix is not diagonal (denoted Σ_ϵ

- From the theory we have SMA representation of the form

$$Y_t = \Gamma(L)v_t$$

where $\Gamma(L) = \Gamma_0 + \Gamma_1 L + ...$, u_t iid and the variance covariance of u_t is diagonal (often normalize to identity)

- How to get $\Gamma(L)$ from $C(L)$ and Σ_ϵ? Problem, more unknowns and estimated parameters.

Structural moving average representation of model solution

- Note thatthe starting point of SVAR is that a given model implies SMA
- When a model has a state space representation, then it generally has a SMA representation, that is, from the model you can write out how each endoegenous variable responds to exogenous shocks.
- Example in RBC case

$$C_t = \Pi_{ck}\Pi_{kA}\sum_{i}^{\infty}(\Pi_{kk})^i\sum_{i}^{\infty}\rho^i u_t + \Pi_{cA}\sum_{i}^{\infty}\rho^i u_t$$

- when $K_{t+1} = \Pi_{KK}K_t + \Pi_{kA}A_t$,
- $A_t = \rho A_{t-1} + u_t$
- $c_t = \Pi_{ck}K_t + \Pi_{cA}A_t$

Main identification issue

- Can rewrite SMA in from of reduced form representation

$$Y_t = \Gamma(L)\Gamma_0^{-1}\Gamma_0 v_t$$

- Note that this is equivalent to wold representation with the mapping

$$\epsilon_t = \Gamma_0 v_t$$

$$C(L) = \Gamma(L)\Gamma_0^{-1}$$

$$\Gamma_0\Gamma_0' = \Sigma_\epsilon$$

Main identification issue

- Can we find a solution to this system of equations

$$C(L) = \Gamma(L)\Gamma_0^{-1}$$

$$\Gamma_0\Gamma_0' = \Sigma_\epsilon$$

- Not with any restrictions, as this has $\frac{n(n-1)}{2}$ to many parameters (n number of variables and shocks)
- So need to use "theory" to impose some restrictions. Different approaches: impact (short run) restrictions, long run restrictions, sign restrictions, variance restrictions.....Full structural corresponds to cross-equation restrictions
- Start with impact restrictions

Short Run restrictions

- Use theory to place restrictions of Γ_0
- Need enough restrictions to solve $\Gamma_0\Gamma_0' = \Sigma_\epsilon$
- Generally set as zero restrictions
- Sometimes in triangular form (ordering matters). Choleski equivalent.
- Bloc recursive structure can be used to identify one shocks.

Review of identification issue

- Much of the debate in Svar literature has nothing to do with dynamics. It can be well understood in static setup

Identification

- Consider model of the form (under identified):

$$y_t^d = -\alpha_1 P_t + \epsilon_{1t}$$

$$Y_t^s = \beta_1 P_t + \epsilon_{2t}$$

$$Y_t^d = Y_t^s$$

- Solving implies

$$P_t = \frac{1}{\beta + \alpha}(\epsilon_{1t} - \epsilon_{2t})$$

$$Y_t = \frac{\beta}{\beta + \alpha}(\epsilon_{1t}) + \frac{\alpha}{\beta + \alpha}(\epsilon_{2t})$$

- Since the variance co-variance has only three elements, and model has four parameters, un-identified.

Standard Identification: IV

- Alternative (just identified)

$$y_t^d = -\alpha_1 P_t + \alpha_2 Z_{1t} + \epsilon_{1t}$$

$$Y_t^s = \beta_1 P_t + \beta_2 Z_{2t} + \epsilon_{2t}$$

with $Cor(Z_{it}, \epsilon_{it}) = 0$,

- Over-identified

$$y_t^d = -\alpha_1 P_t + \alpha_2 Z_{1t} + \alpha_3 X_{1t} + \epsilon_{1t}$$

$$Y_t^s = \beta_1 P_t + \beta_2 Z_{2t} + \beta_3 X_{2t} + \epsilon_{2t}$$

with $Cor(Z_{it}, \epsilon_{it}) = 0$, and $Cor(X_{it}, \epsilon_{it}) = 0$

Impact restriction identification

- If no additional observables: for just identification.

$$Y_t^d = \epsilon_{1t}$$

$$Y_t^s = \beta_1 P_t + \epsilon_{2t}$$

with $Cor(\epsilon_{1t}, \epsilon_{2t}) = 0$

Other sources of identification

- Models with rational expectations often imply cross-equations restrictions which can help with identification. But these depend tightly on the theory, and are not exploited generally in Svar literature.
- Example:

$$Y_t = \alpha E_t[Y_{t+1}] + X_t$$

$$X_t = \rho_1 Z_t + \rho_2 Z_{t-1} + \rho_3 Z_{t-2}$$

where Z_t iid. Implies

$$Y_t = (\rho_1 + \alpha\rho_2 + \alpha^2\rho_3))Z_t + (\rho_2 + \alpha\rho_3)Z_{t-1} + \rho_3 Z_{t-1}$$

- with X_t being an MA(3), model over-identified.

Example

- Identify effects of monetary shocks
- Ref: Christiano, Eichenbaum and Evans (Handbook and JPE 2005)
- Only getting one shock. Bloc recursive idea.

Long Run restrictions

- Recall of main issue in Vars
- Suppose we have in mind an environment where the data is generated by a model of the form:

$$Y_t = \Gamma(L)u_t$$

where Y_t is a vector of variable explained by the model, $\Gamma(L) = \Gamma_0 + \Gamma_1 L + ... + \Gamma_n L^n$ is a polynomial in the lag operator, and u_t is a vector of iid structural shocks, with the variance co-variance matrix of u_t being diagonal (choose identity by nomalization). For data on Y_t, we can estimate a VAR and retrieve a reduced form moving average representation, of the form

$$Y_t = C(L)\epsilon_t$$

where $C(L) = I + C_1 L + ... + C_n L^n$, and the variance co-variance matrix of ϵ is Σ.

Long Run restrictions

- Note the following identities.

$$C(L) = \Gamma(L)\Gamma_0^{-1}, \qquad \Gamma_0 U_t = \epsilon_t$$

$$\Gamma_0 \Gamma_0' = \Sigma$$

Long run restrictions

- Identification issue: can we recover $\Gamma(L)$ from $C(L)$ and Σ.
- In general, answer is not. $\frac{n(n-1)}{2}$ more unknows that observables.
- If environment suggests some restrictions on Γ, then maybe possible.
- Impact restriction: direct restrictions on Γ_0, long run restrictions $\sum_{i=1}^{n} \Gamma_i$ (possible other cross-equations restrictions, but not commonly used).
- If some elements of Y_t are first differences, then $\Gamma(1)$ correspond to the long run effect of a shock on the level.
- Idea is to solve

$$\Gamma_0 \Gamma_0' = \Sigma$$

using restriction

$$C(1)\Gamma_0 = \Gamma(1)$$

- See other file.
- Examples: Gali (AER 1999), Blanchard Quah (AER 89).

Assignment

- Get J-multi (free program for doing long-run and short run structural vars.
- Using data on Y/h and h, and TFP (from John Fernald's site), use long run and short run restrictions to identify effects of TFP shocks.

More Alternatives

- Mixing short and long run (see Beaudry-Lucke)
- Sign restrictions
- Max varainces and Barsky Sims method
- FAvars

Issues

- Co-integration: Singularity of Wold representation. Error Correction formulation. Estimating in Levels
- Fundamental representation and invertibility

Summary

- Estimation and evaluation of macro models can be approached as system estimation amd based on state space.
- However, not always informative or asking too much from a model
- May want to use model to describe how economy reacts to one shock. This suggest the use of SVAR methods.
- SVAR methods can be used to either evaluate model or estimate them, or both.

Vocabulary review

- The structural impulse response (or structural moving average representation) corresponds to the impulse response associated with theoretically meaningful shocks. In general, the structural moving average representation correspond to the moving average derived from a theory where the impulse are given names.
- The reduced form moving average, which can inferred from data (some issues), corresponds to a representation of the form

$$Y_t = C(L)\hat{\mu}_t$$

where the first matrix in $C(L)$ is the identity matrix, and the reduced form errors can be thought as linear combinations of structural errors. Accordingly, the variance co-variance matrix of the reduced form errors is not diagonal (denote by Σ).

Vocabulary review

- From a structural model, we can derive its structural moving average representation (say denoted $\Gamma(L)\mu_t$), and we can also derive its reduced form moving average by application of

$$C(L) = \Gamma(L)(\Gamma^0)^{-1}$$

,
with $\Sigma = \Gamma^0(\Gamma^0)'$

- The identification issue is about inferring $\Gamma(L)$ from $C(L)$ and Σ
- A Choleski decomposition of a symmetric matrix, say Σ, corresponds to finding the triangular matrix K such that

$$KK' = \Sigma$$

- We can always get an orthogonalized moving average representation by using a Choleski decomposition. However, that does not mean we has isolated the structural moving average representation. This only happens if the model admits the right impact restriction. (ordering matters)

Alternative formulation of identification problem

- Can always write

$$\Sigma_\epsilon = PQQ'P$$

- Where P P' is an orthogonalization of Σ and Q is an orthonormal matrix
- with

$$\epsilon_t = PQv_t$$

- So identification problem reduces to find Q.