Imposing structural identifying restrictions in GMA models

Regis Barnichon
CREI, Universitat Pompeu Fabra, CEPR

Christian Matthes
Federal Reserve Bank of Richmond

February 2016

1 Introduction

Although little discussed in Barnichon and Matthes (2016), an important advantage of GMAs is the ease with which one can impose structural identifying assumptions. Since the coefficients of the Gaussian mixtures are directly interpretable in terms of impulse response functions, imposing restrictions on the GMA coefficients amounts to directly imposing restrictions on the shape of the impulse response function. This is in contrast with standard VAR approaches, where imposing constraints on the shape of the impulse responses can be difficult, because the impulse responses are (non-trivial) non-linear transformations of the VAR parameters. In this note, we describe how to implement the main structural identifying assumptions used in the literature –(i) identification from short-run or long-run restrictions, and (ii) sign restrictions–, but also how to impose more general forms of identifying restrictions that cannot be easily imposed in traditional VAR settings. We describe the implementation of the restriction in linear GMA models, but an implementation in more general non-linear models is also possible, as shown in Barnichon and Matthes (2016) for the cases of recursive short-run identification restrictions or sign restrictions.

As in Barnichon and Matthes (2016), we consider the structural moving-average model

$$\mathbf{Y}_t = \sum_{k=0}^{K} \mathbf{\Psi}_k \boldsymbol{\varepsilon}_{t-k} \tag{1}$$

with $\boldsymbol{\varepsilon}_t = (.., \varepsilon_{jt}, ..)'$, $i \in \{1, .., M\}$, the vector of structural shocks and $\mathbf{\Psi}_k = \left(\psi_{k,ij}\right)$ an $(M \times M)$ matrix. For $k > 0$, the impulse response functions are modeled as mixtures of Gaussian basis functions so that $\psi_{k,ij}$, the impulse response of variable i to shock ε_j at horizon k, is given by

$$\psi_{k,ij} = \sum_{n=1}^{N} a_{ij,n} e^{-\left(\frac{k-b_{ij,n}}{c_{ij,n}}\right)^2}$$

with $\{a_{ij,n}, b_{ij,n}, c_{ij,n}\}$ parameters to be estimated.

2 Common identification schemes

All identifying schemes impose some restrictions on the parameters to be estimated. To implement parameter restrictions, we assign a minus infinity value to the likelihood whenever the restrictions are not met.[1] We now show how to implement popular identification schemes.

2.1 Short-run restrictions

In a just-identified model, short-run restrictions define $\frac{M(M-1)}{2}$ restrictions on $\mathbf{\Psi}_0$ with $\mathbf{\Psi}_0$ the contemporaneous impact matrix of size $(M \times M)$. A common restriction on $\mathbf{\Psi}_0$ is that $\mathbf{\Psi}_0$ is lower triangular. However, other restrictions on $\mathbf{\Psi}_0$ are possible as long as $\mathbf{\Psi}_0$ is invertible. Indeed, and as described in Barnichon and Matthes (2016), the only requirement to recursively construct the likelihood at time t, is that that the system of equations from (1)

$$\mathbf{\Psi}_0 \boldsymbol{\varepsilon}_t = \mathbf{u}_t \tag{2}$$

where $\mathbf{u}_t = \mathbf{Y}_t - \sum_{k=0}^{K} \mathbf{\Psi}_k \boldsymbol{\varepsilon}_{t-1-k}$ has a unique solution vector $\boldsymbol{\varepsilon}_t$. That is, that the shock vector $\boldsymbol{\varepsilon}_t$ is uniquely determined given a set of model parameters and the history of variables up to time t. This is ensured by having $\mathbf{\Psi}_0$ is invertible.[2]

For instance, in a partial recursive identification scheme, one can posit $\mathbf{\Psi}_0$ then has its last column filled with 0 except for the diagonal coefficient, that is that the shock of interest (ordered last) has no no contemporaneous effect on the other variables (this is assumption used in Barnichon and Matthes, 2016). Then, the only restriction on $\mathbf{\Psi}_0$ necessary to construct the likelihood is that the submatrix $\tilde{\mathbf{\Psi}}_0$ made of the first $(M-1)$ rows and $(M-1)$ columns of $\mathbf{\Psi}_0$ is invertible.

2.2 Long-run restrictions

Long-run restrictions are often used in two-variable VARs where one of the variables is entered in first-difference (e.g., Blanchard and Quah 1989, Gali, 1999). A popular example is for instance

$$Y_t = \begin{pmatrix} d\ln\left(\frac{y_t}{h_t}\right) \\ U_t \end{pmatrix}, \varepsilon_t = \begin{pmatrix} \varepsilon_t^a \\ \varepsilon_t^u \end{pmatrix}$$

[1] Equivalently, in the MCMC stage, we reject all draws that do not satisfy the parameter restrictions.
[2] This is ensured by assigning a minus infinity value to the likelihood whenever $\mathbf{\Psi}_0$ is not invertible.

where $\frac{y_t}{h_t}$ is output per hours worked and U_t is the unemployment rate. ε_t^a is a technology shock, and ε_t^u is a non-technology shock.

The identification assumption is that non-technology shocks have no long-run effect on productivity. This means that

$$\sum_{k=0}^{K} \psi_{k,21} = 0$$

or

$$\psi_{0,21} = -\sum_{k=1}^{K} \psi_{k,21}.$$

This restriction can be easily implemented: we can draw the $\{a_{n,21}, b_{n,21}, c_{n,21}\}$ GMA parameters for $\{\psi_{k,21}\}_{k>0}$ and impose that $\psi_{0,21} = -\sum_{k=1}^{K} \psi_{k,21}$. The other parameters from $\mathbf{\Psi}_0$ being drawn in the usual fashion, only that one must discard draws for which $\mathbf{\Psi}_0$ is non-invertible.

2.3 Sign restrictions

Since the coefficients of the Gaussian mixtures are directly interpretable in terms of impulse response functions, imposing sign restrictions is very simple in GMA models, whether we want to impose sign-restrictions on the impact coefficients (captured by $\mathbf{\Psi}_0$) and/or sign restrictions on the post-impact coefficients $\mathbf{\Psi}_{k,k>0}$. Since $\mathbf{\Psi}_{k,k>0}$ is determined by the $\{a_n, b_n, c_n\}$ GMA coefficients, one can impose sign restrictions by imposing sign restrictions on the loading of the different Gaussian basis functions, that is by imposing sign restrictions on the $\{a_n\}_{n=1}^N$ coefficients.

More generally, and in line with the insights from Baumeister and Hamilton (2015), imposing sign-restrictions would take the form of priors on the coefficients of $\mathbf{\Psi}_0$ or on the $\{a_n\}_{n=1}^N$ coefficients in a GMA(N) model.

2.4 General identification schemes with a Bayesian formulation

More generally, because GMAs work directly with the structural moving-average representation, the parameters to be estimated can be easily interpreted as "features" of the impulse responses and many set identification schemes can be easily implemented. Using the insights from Baumeister and Hamilton (2015), one can (in addition to possible sign restrictions) posit priors on the *shape* of the impulse responses, posit priors on the location of the peak effect, posit priors on the persistence of the effect of the shock, among other possibilities.

For instance, in a GMA(1) model, one could impose that the impulse response function is monotonically decreasing by imposing $b \leq 0$. Alternatively, if one believed that the peak effect of a shock occurred between two and six quarters after a shock, one could impose that b has a prior centered at 4 quarters with 90 percent of the mass between 2 and 6 quarters. If would believed that the effect of a particular shock dies out rapidly, one could impose a prior for c centered around a low value,[3] etc...

In higher-order GMAs, similar interpretations apply to the different Gaussians used to approximate the impulse response. Although deserving a much more thorough study outside the scope of this short note, in a GMA(N) where the N Gaussian basis functions are chosen to being approximately orthogonal (i.e., their inner product is close to zero), one can interpret an impulse response function as being *decomposed* into a sum of N (approximately) independent effects of the shock.[4] With a GMA(N), the a_n, b_n and c_n coefficients of the nth Gaussian basis function ($n \in \{1, ..., N\}$) are then respectively the magnitude, location and persistence of the "nth effect" of the shock. One could achieve set identification by imposing priors on the $\{a_n, b_n, c_n\}$ coefficients. For instance, if a shock leads to an oscillating pattern, the N Gaussian basis functions would capture the different waves of the impulse response, and one could impose restrictions of the shape (e.g., location and duration) of each of these waves.

[3] Recall that the amount of time τ required for the effect of a shock to be 50% of its maximum value is given by $\tau = c\sqrt{\ln 2}$.

[4] This interpretation is similar to B-splines smoothing, when one projects a function of interest on a small set of approximately orthogonal B-splines (see Tibshirani et al., 2009).

References

[1] Barnichon R. and C. Matthes. "Gaussian Mixture Approximations of Impulse Responses and the Non-Linear Effects of Monetary Shocks," Working Paper, 2015

[2] Baumeister, C. and J. Hamilton, "Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information," Econometrica, 83(5), 1963-1999, 2015.

[3] Blanchard, O. and D. Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," American Economic Review, 79(4), pages 655-73, September 1989.

[4] Gali, J. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," American Economic Review, vol. 89(1), pages 249-271, March 1999.

[5] Hastie, T., R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer 2009.