Structural Analysis with Classical and Bayesian Large Reduced Rank VARs

Andrea Carriero *George Kapetanios*
Queen Mary, University of London
Massimiliano Marcellino
EUI, Bocconi and CEPR

National Bank of Serbia, 19 October 2012

Introduction

- Econometric models for large datasets widely used in recent applied econometrics literature
- Two main approaches: factor models and VARs

Factor models

- Small scale factor models. Geweke (1977), Sargent and Sims (1977).
- Factor models for large datasets started classical and mostly non-parametric (Forni, Hallin, Lippi, and Reichlin (2000), Stock and Watson (2002))
- Then turned parametric (Structural FAVAR) Kose, Otrock, and Whiteman (2003), Bernanke, Boivin, and Eliasz (2005), Doz, Giannone, and Reichlin (2006).
- Problems:
  - often two step approach (estimate factors, then treat them as known, though full ML possible, e.g. Doz, Giannone, and Reichlin (2006))
  - relies on N diverging for consistent estimation
  - complex to test hypotheses and/or impose restrictions on factors
  - unclear why VAR for factors (Dufour and Stevanovic (2010) - FAVARMA),

BVARs

- Recently, large Bayesian VARs proposed as alternative to factor models, Banbura, Giannone, Reichlin (2010).
- Large classical VARs not feasible, unless constraints imposed: reduced rank VARs, Carriero, Kapetanios, Marcellino (2011).
- Reduced rank VARs also similar to factor models $\longrightarrow$ focus on how to use RR-VARs for structural analysis

RR-(B)VARs

- We suggest to use a model that bridges BVARs and factor models.
- Specifically, we propose to impose reduced rank restrictions in a BVAR that, as we will see, makes it similar to a factor model in terms of having a smaller set of key shocks or variables, but preserves the attractive features of a BVAR, substantially reducing its parameter dimensionality.
- Reduced rank BVARs (RR-BVARs) have been previously considered in the literature, see e.g. Geweke (1996) and Carriero Kapetanios Marcellino (2011) in a small and large datasets context, respectively, but not for structural analysis.
- Our main contribution is to show how a RR-BVAR can be used to compute the response functions to structural shocks, for which we introduce and study the properties of (both classical and Bayesian) methods for estimation, inference and rank determination.

Structure of presentation

- Specification
- Estimation and Rank determination (classical and Bayesian)
- Empirical analysis 1: shocks to factors - exploiting factor structure
- Empirical analysis 2: monetary policy shock - traditional VAR type analysis

RR-VAR

- Consider the N-vector $Y_t = (y_{1,t}, y_{2,t}, ..., y_{N,t})'$:

$$Y_t = \Phi(L) Y_t + \epsilon_t, \tag{1}$$

  where $\Phi(L) = \Phi_1 L + ... + \Phi_p L^p$ and ϵ_t are i.i.d. $N(0, \Sigma)$.

- Assume $\Phi(L) = A(L)B(L)$, where $A(L) = A_1 L + ... + A_{p_1} L^{p_1}$, each A_i is $N \times r$, $B(L) = B_0 + B_1 L + ... + B_{p_2} L^{p_2}$ and each B_i is $r \times N$, with $p_1 + p_2 = p$, $p_1 \geq 1$, $p_2 \geq 0$. Then

$$Y_t = A(L)B(L)Y_t + \epsilon_t = \sum_{u=1}^{p_1} \sum_{v=0}^{p_2} A_u B_v Y_{t-u-v} + \epsilon_t \tag{2}$$

- If r much smaller than N, RR-VAR has much fewer parameters than VAR. For example, if $N = 50$, $p = 2$ and $r = 2$, there are $N(Np - r(p+1)) = 4700$ parameters less in RR-VAR than VAR (300 vs 5000).

RR-VAR and factors

- In $Y_t = A(L)B(L)Y_t$ the r-dimensional vector of variables:

$$F_t = B(L)Y_t = B_0 Y_t + B_1 Y_{t-1} + ... + B_{p_2} Y_{t-p_2} \tag{3}$$

  can be interpreted as r common factors.
- Indeed, RR-VAR can be written as:

$$Y_t = A(L)F_t + \epsilon_t = \sum_{u=1}^{p_1} A_u F_{t-u} + \epsilon_t, \tag{4}$$

- If $p_2 = 0$, we get:

$$Y_t = \sum_{u=1}^{p} A_u F_{t-u} + \epsilon_t, \tag{5}$$

$$F_t = B_0 Y_t \tag{6}$$

- As in factor models, "loadings" A_u and factor weights B_0 are not uniquely identified, we assume that $B_0 = (I_r, \ \widetilde{B}_0)$.

RR-VAR and factors

- The factors $F_t = B_0 Y_t$ have closed form $VAR(p)$ representation:

$$F_t = B_0 \sum_{u=1}^{p} A_u F_{t-u} + B_0 \epsilon_t = C(L) F_t + u_t \tag{7}$$

with $C(L) = B_0 A(L) = B_0 A_1 L + B_0 A_2 L^2 + ... + B_0 A_p L^p$ and $u_t = B_0 \epsilon_t$ is i.i.d. $N(0, \Omega)$ with $\Omega = B_0 \Sigma B_0'$.

- We can then group together (5) and (7) to form the system

$$\begin{cases} Y_t = A(L) F_t + \epsilon \\ F_t = C(L) F_t + u_t \end{cases} \tag{8}$$

MA representation (1)

- The factors have the following MA representation:

$$F_t = (I - C(L))^{-1} u_t = (I - B_0 A(L))^{-1} B_0 \epsilon_t \tag{9}$$

- The moving average representation associated with (8) is:

$$Y_t = A(L)F_t + \epsilon_t = (A(L)(I - B_0 A(L))^{-1} B_0 + I)\epsilon_t. \tag{10}$$

- A second moving average representation is particularly convenient for structural analysis.

MA representation (2)

- The matrix $B_{0\perp}$ is $(N-r) \times N$ full row rank matrix orthogonal to B_0, i.e. $B_0 B'_{0\perp} = 0$, with rank of $(B'_0, B'_{0\perp})$ equal to N. $B_0 B'_0$ has full rank (as B_0 has full rank) and we have:

$$B'_0(B_0 B'_0)^{-1} B_0 + B'_{0\perp}(B_{0\perp} B'_{0\perp})^{-1} B_{0\perp} = I_N. \tag{11}$$

- Inserting this into the Wold representation (10) yields:

$$Y_t = (B'_0(B_0 B'_0)^{-1} + A(L)(I - B_0 A(L))^{-1}) B_0 \epsilon_t + B'_{0\perp}(B_{0\perp} B'_{0\perp})^{-1} B_{0\perp} \epsilon_t, \tag{12}$$

- Since $B_0 \epsilon_t = u_t$, where $F_t = C(L) F_t + u_t$, we have:

$$Y_t = (B'_0(B_0 B'_0)^{-1} + A(L)(I - B_0 A(L))^{-1}) u_t + B'_{0\perp}(B_{0\perp} B'_{0\perp}) B_{0\perp} \epsilon_t. \tag{13}$$

- So, each element of Y_t is driven by a set of r common errors u_t -that are the drivers of the factors F_t- and by elements of $B_{0\perp}\epsilon_t$, where u_t and $B_{0\perp}\epsilon_t$ are orthogonal.

MA representations

- It is then possible to have the following moving average representations:

$$Y_t = (A(L)(I - B_0 A(L))^{-1} B_0 + I)\epsilon_t \tag{14}$$

and:

$$Y_t = (B_0'(B_0 B_0')^{-1} + A(L)(I - B_0 A(L))^{-1})u_t + B_{0\perp}'(B_{0\perp} B_{0\perp}')B_{0\perp}\epsilon_t. \tag{15}$$

- Representation (14) is similar to the one used in the BVAR literature. There are as many shocks as variables (N)
- Representation (15) is similar to the one used in the factors literature. There is a reduced number of shocks (r) which drive all the factors, which in turn drive all the variables in Y_t.
- We do not prefer either representation, we suggest to use the one that is more suited to address the specific empirical problem under analysis.
- There is a case where shocking the factors or shocking the variables produces the same responses and this is of course when the factors are equal to a subset of the variables and we shock one of the variables in this subset.

MA representations and impulse responses

- The structural shocks v_t driving F_t are obtained from the reduced form errors u_t using any technique adopted in the structural VAR and structural FAVAR literatures, e.g. Bernanke et al. (2005) or Eickmeier et al. (2009).
- Simplest option is the Choleski-type decomposition $\Omega = P^{-1}SP^{-1'}$, where P is lower triangular and S is diagonal. Hence,

$$v_t = Pu_t. \tag{16}$$

- Combining this with (15) yields:

$$Y_t = (B_0'(B_0 B_0')^{-1} + A(L)(I - B_0 A(L))^{-1})P^{-1}v_t + B_{0\perp}'(B_{0\perp}B_{0\perp}')B_{0\perp}\epsilon_t \tag{17}$$

from which impulse response functions can be easily computed.

- For the alternative MA representation, the impulse responses can be based on:

$$Y_t = (A(L)(I - B_0 A(L))^{-1}B_0 + I)\Lambda^{-1}\epsilon_t^* \tag{18}$$

where ϵ_t^* are structural shocks and Λ^{-1} is such that $\Sigma = \Lambda^{-1}R\Lambda^{-1\prime}$, with R diagonal.

RR-VAR and factor models

- In summary, RR-VAR is similar to the generalized dynamic factor model of Forni, Hallin, Lippi, and Reichlin (2001) and Stock and Watson (2002a, 2002b), and even more to the parametric versions of these models in the FAVAR literature, e.g. Bernanke et al. (2005) and Doz, Giannone, and Reichlin (2006).
- But also possibly important differences

RR-VAR and factor models

- Testing/imposing specific restrictions on the factors F_t, such as equality of one factor to a specific economic variable, is much simpler in the $RR - VAR$ context (restrictions on B_0) than in a factor context (see Bai and Ng 2006, 2010, Chahrour 2011). This greatly helps in structural interpretation.
- In the Bayesian implementation, one can also impose priors on what variables "go" into which factors by putting an informative prior on B_0.
- In the factor literature factors are unobservable and can be consistently estimated only when N diverges. Within an $RR - VAR$ context it is possible to consistently estimate the "factors" F_t even when N is finite.
- In the factor literature precise conditions on common and idiosyncratic components are needed. In $RR - VAR$ one could constrain the var cov matrix of $B_{0\perp}\epsilon_t$ in (13), or $A(L)$ in (8). But these are optional.
- Factors models estimated by PC do not necessarily have an exact VAR representation (Dufour and Stevanovic (2010)), while this is the case within the $RR - VAR$ context.

Estimation

- For estimation, we compactly rewrite the RR-VAR as in Reinsel (1983):

$$Y_t = AZ_{t-1} + \epsilon_t, \tag{19}$$

where
$Z'_{t-1} = (F'_{t-1} ..., F'_{t-p}) = (Y'_{t-1} B'_{0,} ..., Y'_{t-p} B'_{0,}) = (Y'_{t-1}, ..., Y'_{t-p})(I_p \otimes B'_0)$ and $A = (A_1, ..., A_p)$,with $B_0 = (I_r,\ \widetilde{B}_0)$. Defining $Y = (Y_1, ..., Y_T)'$ and $Z = (Z_0, Z_1, ..., Z_{T-1})'$ and $E = (\epsilon_1, ..., \epsilon_T)'$

- Stacking the equations in (19) for $t = 1, ..., T$ we have

$$Y = ZA + E, \tag{20}$$

where $VAR(E) = \Omega = (\Sigma \otimes I_T)$.

Estimation via Maximum Likelihood

- Reinsel (1983) did most of the work. He provides the FOCs and updating rule for the gradient of the ML estimator for this case.
- ML estimates can be obtained by iterating over the first order conditions of the maximization problem. This does not involve numerical optimization if B_0 is left free of restrictions.
- The likelihood function is:

$$-0.5T\log|\Omega| - 0.5\Sigma_{t=1}^{T}(Y_t - AZ_{t-1})'\Omega^{-1}(Y_t - AZ_{t-1}) \tag{21}$$

- For any A and $\tilde{B}_0$ the maximization with respect to Ω yields:

$$\hat{\Omega} = (Y - ZA')'(Y - ZA')/T \tag{22}$$

Estimation via Maximum Likelihood

- The partial derivatives with respect to A (given $\tilde{B}_0$ and Ω):

$$\frac{\partial l}{\partial vec(A')} = \Sigma_{t=1}^{T}(I_N \otimes Z_{t-1})'\Omega^{-1}(I_N \otimes Z_{t-1}) = 0 \tag{23}$$

- The partial derivatives with respect to $\tilde{B}_0$:

$$\frac{\partial l}{\partial vec(\tilde{B}_0)} = \Sigma_{t=1}^{T} U_{t-1} A'\Omega^{-1}\{Y_t - (I_N \otimes Z'_{t-1})\alpha\} = 0 \tag{24}$$

  where $U_{t-1} = (I_r \otimes Y_{2,t-1}, ..., I_r \otimes Y_{2,t-p})$ and $Y'_{2,t}$ comes from partitioning Y'_t in the first r and last $N-r$ components: $Y'_t = (Y'_{1,t}, Y'_{2,t})$.

- Reinsel suggested to solve in turn equations (22), (23) and (24) until convergence is achieved, and established consistency and asymptotic normality of this estimator.
- It is also possible to impose constraints on $\tilde{B}_0$ but then the iterative scheme described above is no longer available.

Estimation via Markov Chain Monte Carlo

- We derive the conditional distributions and provide a new MCMC algorithm for the estimation of

$$Y_t = AZ_{t-1} + \epsilon_t, \tag{25}$$

  where
  $Z'_{t-1} = (F'_{t-1} ..., F'_{t-p}) = (Y'_{t-1}B'_{0,} ..., Y'_{t-p}B'_{0,}) = (Y'_{t-1}, ..., Y'_{t-p})(I_p \otimes B'_0)$ and $A = (A_1, ..., A_p)$, with $B_0 = (I_r,\ \widetilde{B}_0)$.
- The model contains three sets of parameters, in the matrices A, $\widetilde{B}_0$, and Σ. The joint posterior distribution $p(A, \widetilde{B}_0, \Sigma | Y)$ has not a known form, but it can be simulated by drawing in turn from the conditional posterior distributions $p(A|\widetilde{B}_0, \Sigma, Y)$, $p(\widetilde{B}_0|A, \Sigma, Y)$, and $p(\Sigma|A, \widetilde{B}_0, Y)$.

Priors

- Assume a Normal-Inverse Wishart prior for A and Σ :

$$A|\Sigma \sim N(A_0, \Sigma \otimes \Omega_0),\ \Sigma \sim IW(S_0, v_0). \tag{26}$$

with:

$$\Omega_0 = \tau I \tag{27}$$

$$A_0 = 0 \tag{28}$$

$$S_0 = S_{AR} \tag{29}$$

$$v_0 = N + 2 \tag{30}$$

where S_{AR} is a diagonal matrix of residual sum of squares from univariate regressions and where $\sqrt{\tau} = 0.05$.

- We assume both a flat or an informative prior on $\widetilde{B}_0$. The informative prior would be centered at the PC estimate, with s.d. 0.1.

Posteriors

- Under the knowledge of $\widetilde{B}_0$ and Y the variable Z_{t-1} is known, and (25) is a simple multivariate regression model as in Zellner (1973). Then the conditional posterior distributions are:

$$A|\Sigma, \widetilde{B}_0, Y \sim N(\bar{A}, \Sigma \otimes \bar{\Omega}), \ \Sigma|\widetilde{B}_0, Y \sim IW(\bar{S}, \bar{v}). \tag{31}$$

with:

$$\bar{\Omega} = (\Omega_0^{-1} + Z'Z)^{-1} \tag{32}$$

$$\bar{A} = \bar{\Omega}(\Omega_0^{-1} A_0 + Z'Y) \tag{33}$$

$$\bar{S} = S_0 + Y'Y + A_0'\Omega_0^{-1} A_0 - \bar{A}'\bar{\Omega}^{-1}\bar{A} \tag{34}$$

$$\bar{v} = v_0 + T \tag{35}$$

Posterior simulation

- Draws from $p(A, \Sigma | \widetilde{B}_0, Y)$ can be easily obtained by MC integration by generating a sequence of m draws $\{\Sigma_j\}_{j=1}^{m}$ from $\Sigma | \widetilde{B}_0, Y \sim IW(\bar{S}, \bar{v})$ and then for each j drawing from $A | \widetilde{B}_0, \Sigma, Y \sim N(\bar{A}, \Sigma_j \otimes \bar{\Omega})$, which provides the sequence $\{A_j\}_{j=1}^{m}$.
- The prior features a Kronecker structure that restricts somehow the way shrinkage can be imposed, but dramatically improves the computation time.
- In particular the inversion of the matrix $\Omega_0^{-1} + Z'Z$ is not problematic as it is of dimension pr.
- Restrictions on A could also be imposed.

Posterior simulation

- Drawing from $p(\widetilde{B}_0|A, \Sigma, Y)$ is less simple, as B_0 has identification restrictions and enters nonlinearly in the model : $Y_t = A(L)B_0 Y_t$.
- To draw $\widetilde{B}_0|A, \Sigma, Y$ we use a Metropolis step.
- In particular, for each draw in the sequence $\{A_j, \Sigma_j\}_{j=1}^m$ a candidate $\widetilde{B}_{0j}^*$ for each element of $\widetilde{B}_0$ is drawn by sampling from a random walk:

$$vec(\widetilde{B}_{0j}^*) = vec(\widetilde{B}_{0j-1}) + c\eta_t \qquad (36)$$

where η_t is a standard normal i.i.d. process and c is scaling factor calibrated in order to have about 30% rejections. The candidate draw is then accepted with probability:

$$\alpha_k = \min\left\{\frac{p(\widetilde{B}_{0j}^*|\Sigma, \widetilde{B}_0, Y)}{p(\widetilde{B}_{0j-1}|\Sigma, \widetilde{B}_0, Y)}, 1\right\} \qquad (37)$$

If the draw is accepted, then $\widetilde{B}_{0j} = \widetilde{B}_{0j}^*$, otherwise $\widetilde{B}_{0j} = \widetilde{B}_{0j-1}$.

Drawing impulse responses

- Drawing in turn from $p(A|\widetilde{B}_0, \Sigma, Y)$, $p(\Sigma|\widetilde{B}_0, Y)$, and $p(B_0|A, \Sigma, Y)$ provides a sequence of m draws $\left\{A_j, \Sigma_j, \widetilde{B}_0\right\}_{j=1}^{m}$ from the joint posterior distribution of $\widetilde{B}_0, A, \Sigma$.
- Each draw can be then inserted into

$$Y_t = (B_0'(B_0 B_0')^{-1} + A(L)(I - B_0 A(L))^{-1})P^{-1} v_t + B_{0\perp}'(B_{0\perp} B_{0\perp}')B_{0\perp}\epsilon_t, \quad (38)$$

- or

$$Y_t = (A(L)(I - B_0 A(L))^{-1} B_0 + I)\Lambda^{-1}\epsilon_t^* \quad (39)$$

which can be used to derive the impulse response functions for any horizon.

Determining the rank of the system

Classical

- Two main approaches: information criteria or sequential testing.
- Standard info criteria can be used. An attractive feature is that both r and the number of lags can be jointly determined.
- Sequential testing: starting with the null hypothesis of $r = 1$, a sequence of tests is performed. If the null hypothesis is rejected, r is augmented by one and the test is repeated. When the null cannot be rejected, r is adopted as the estimate of the rank of each matrix A_i in (19).
- Tests described in the paper.

Determining the rank of the system

Bayesian

- Compute the marginal data density as a function of the chosen r. Such density is given by:

$$p_r(Y) = \int p_r(Y|\theta)p(\theta)d\theta \tag{40}$$

- The optimal rank for the system is associated with the model featuring the highest data density:

$$r^* = \arg\max_r \; p_r(Y) \tag{41}$$

- Note r^* corresponds to the posterior mode of r under a flat prior.
- The marginal data density $p_r(Y)$ can be approximated numerically by using Geweke's (1999) modified harmonic mean estimator.

Shocks to factors

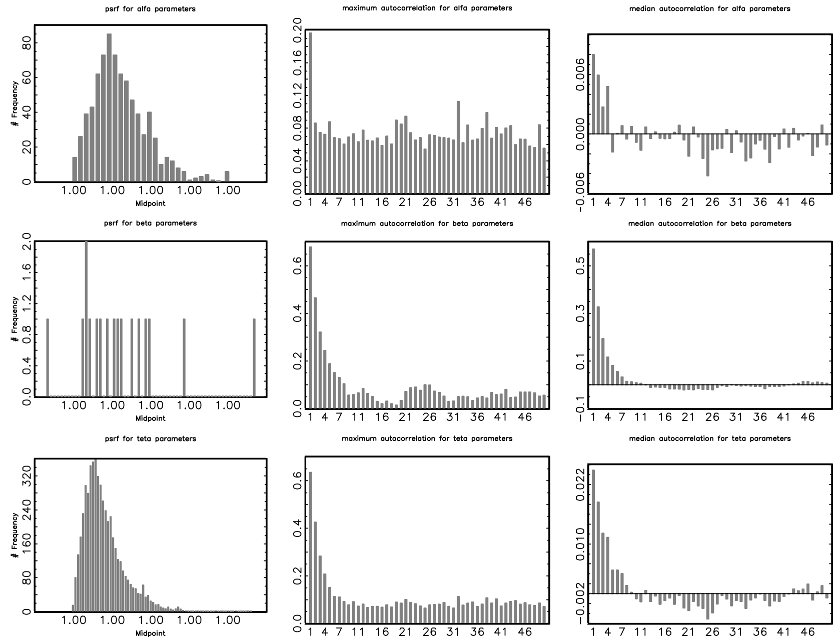
- We analyze the effects of a shock to factors using the "medium" dataset of Banbura, Giannone, Reichlin (2010)
- We set the system rank to 3 and the lag length to 13.
- We identify an output factor, a price factor, and a financial/monetary factor by imposing restrictions on the matrix B_0

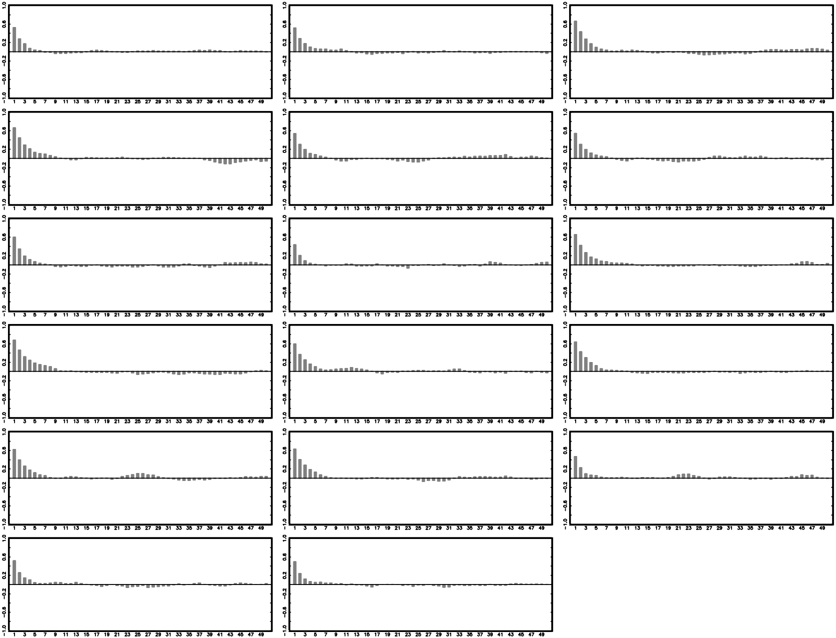
Data

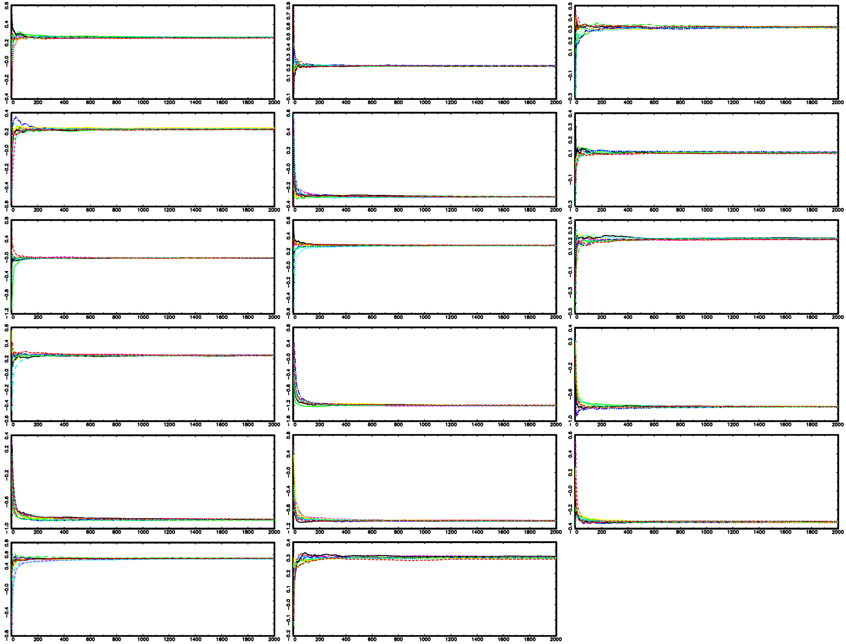
Variable	Short name	F1	F2	F3
Employees on nonfarm payroll	EMP: TOTAL	1	0	0
CPI, all items	CPI-U: ALL	0	1	0
Index of sensitive material prices	SENS MAT'LS PRICE	0	$b_{2,3}$	0
Personal income	PI	$b_{1,4}$	0	0
Real Consumption	CONSUMPTION	$b_{1,5}$	0	0
Industrial Production Index	IP: TOTAL	$b_{1,6}$	0	0
Capacity Utilization	CAP UTIL	$b_{1,7}$	0	0
Unemployment rate	U: ALL	$b_{1,8}$	0	0
Housing starts	HSTARTS: TOTAL	$b_{1,9}$	0	0
Producer Price Index (finished goods)	PPI: FIN GDS	0	$b_{2,10}$	0
Implicit price deflator for personal consumption expenditures	PCE DEFL	0	$b_{2,11}$	0
Average hourly earnings	AHE: GOODS	$b_{1,12}$	0	0
Federal Funds, effective	FEDFUNDS	0	0	1
M1 money stock	M1	0	0	$b_{3,14}$
M2 money stock	M2	0	0	$b_{3,15}$
Total reserves of depository institutions	RESERVES TOT	0	0	$b_{3,16}$
Nonborrowed reserves of depository institutions	RESERVES NONBOR	0	0	$b_{3,17}$
S&P's common stock price index	S&P 500	0	0	$b_{3,18}$
Interest rate n treasury bills, 10 year constant maturity	10 YR T-BOND	0	0	$b_{3,19}$
Effective Echange rate	EX RATE: AVG	0	0	$b_{3,20}$

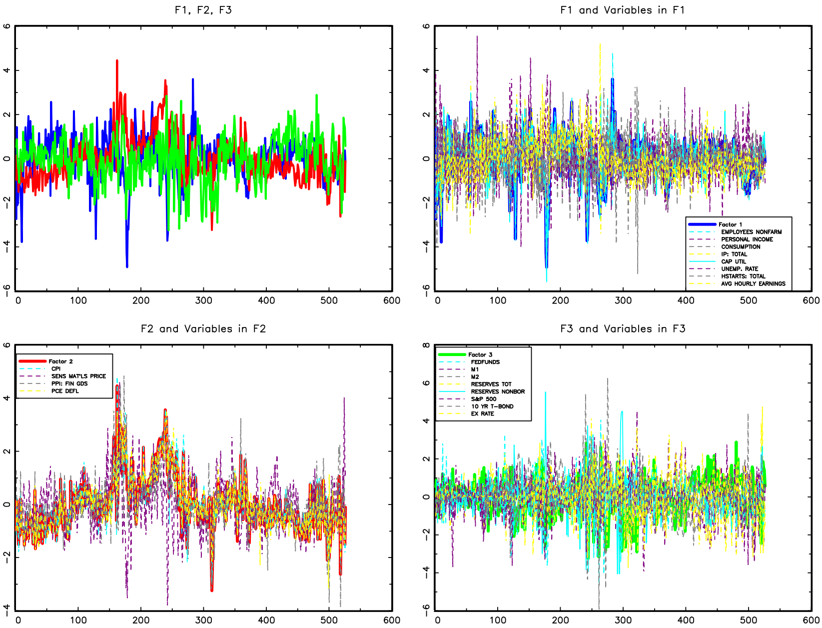
Estimates

	Prior mean (std=0.1)			Posterior mean			Posterior std		
variable	F1	F2	F3	F1	F2	F3	F1	F2	F3
EMP:TOTAL	1	0	0	1	0	0	-	-	-
CPI-U:ALL	0	1	0	0	1	0	-	-	-
SENSMAT PRICE	0	0.27	0	0	0.28	0	-	0.071	-
PI	0.74	0	0	0.30	0	0	0.078	-	-
CONSUMPTION	0.50	0	0	0.24	0	0	0.076	-	-
IP:TOTAL	1.09	0	0	0.38	0	0	0.098	-	-
CAPUTIL	1.08	0	0	0.31	0	0	0.096	-	-
U:ALL	−0.76	0	0	−0.40	0	0	0.083	-	-
HSTARTS	0.23	0	0	0.12	0	0	0.082	-	-
PPI:FINGDS	0	0.89	0	0	0.28	0	-	0.090	-
PCEDEFL	0	0.99	0	0	0.31	0	-	0.090	-
AHE:GOODS	0.06	0	0	−0.006	0	0	0.088	-	-
FEDFUNDS	0	0	1	0	0	1	-	-	-
M1	0	0	−1.80	0	0	−1.31	-	-	0.101
M2	0	0	−1.27	0	0	−0.87	-	-	0.102
RES.TOT	0	0	−1.82	0	0	−1.00	-	-	0.101
RES.NONBOR	0	0	−1.96	0	0	−1.19	-	-	0.105
S&P500	0	0	−0.67	0	0	−0.39	-	-	0.087
10YRT-BOND	0	0	0.59	0	0	0.57	-	-	0.091
EXRATE	0	0	0.46	0	0	0.37	-	-	0.085


psrf for alfa parameters
maximum autocorrelation for alfa parameters
median autocorrelation for alfa parameters
psrf for beta parameters
maximum autocorrelation for beta parameters
median autocorrelation for beta parameters
psrf for teta parameters
maximum autocorrelation for teta parameters
median autocorrelation for teta parameters
Frequency
Midpoint
1.00 1.00 1.00 1.00 1.00
1 4 7 11 16 21 26 31 36 41 46


1.0
0.6
0.2
-0.2
-0.6
-1.0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49




F1, F2, F3
F1 and Variables in F1
Factor 1
EMPLOYEES NONFARM
PERSONAL INCOME
CONSUMPTION
IP: TOTAL
CAP UTIL
UNEMP. RATE
HSTARTS: TOTAL
AVG HOURLY EARNINGS
F2 and Variables in F2
Factor 2
CPI
SENS MAT'LS PRICE
PPI: FIN GDS
PCE DEFL
F3 and Variables in F3
Factor 3
FEDFUNDS
M1
M2
RESERVES TOT
RESERVES NONBOR
S&P 500
10 YR T−BOND
EX RATE

Responses to shocks to factors

- The impulse responses are based on the representation:

$$Y_t = \{B_0'(B_0 B_0')^{-1} + A(L)[I - B_0 A(L)]^{-1}\}P^{-1}v_t + B_{0\perp}'(B_{0\perp}B_{0\perp}')B_{0\perp}\epsilon_t, \qquad (42)$$

  where P^{-1} is the Cholesky factor of the reduced form shocks u_t.
  $(P^{-1}SP^{-1'} = \Omega = B_0\Sigma B')$.

- The s-period ahead response on the factor equation is:

$$\Pi_s = C_1\Pi_{s-1} + ... + C_{\min(s,p)}\Pi_{s-\min(s,p)};\ s > 0 \qquad (43)$$

  with $\Pi_0 = [I - B_0 A(0)]^{-1}P^{-1} = P^{-1}$

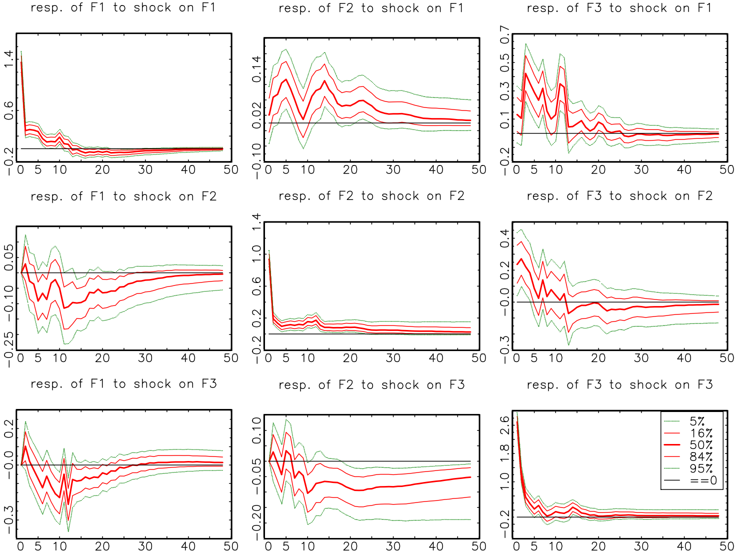
- The s-period ahead response on the VAR equation is:

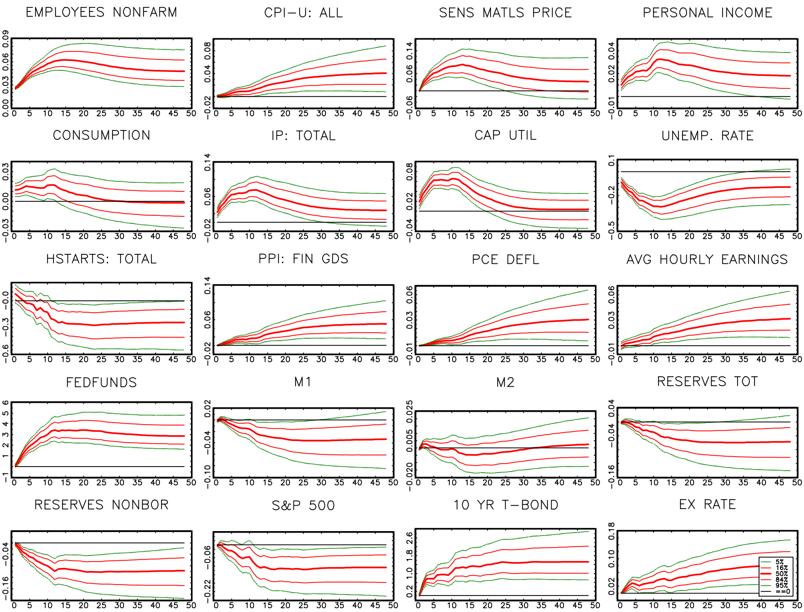
$$\Psi_s = A_1\Pi_{s-1} + ... + A_{\min(s,p)}\Pi_{s-\min(s,p)};\ s > 0 \qquad (44)$$

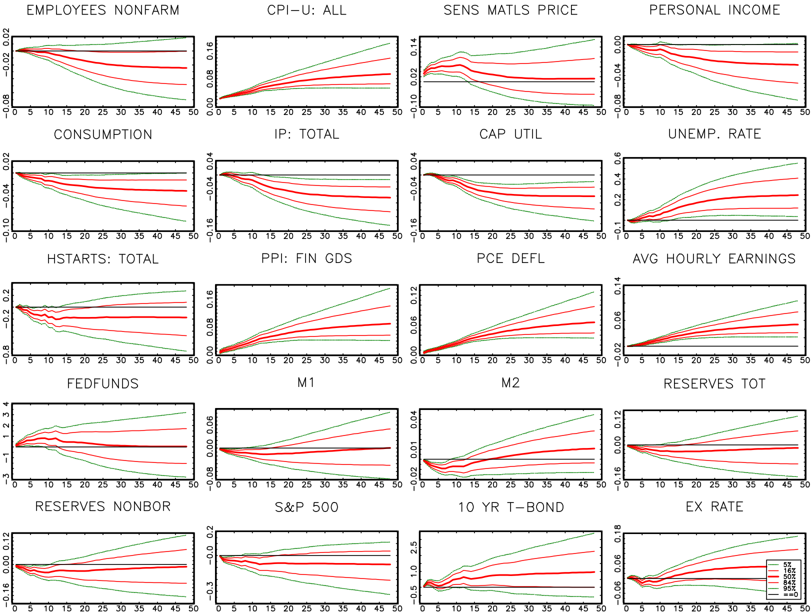
  with $\Psi_0 = \{B_0'(B_0 B_0')^{-1} + A(0)[I - B_0 A(0)]^{-1}\}P^{-1} = B_0'(B_0 B_0')^{-1}P^{-1}$

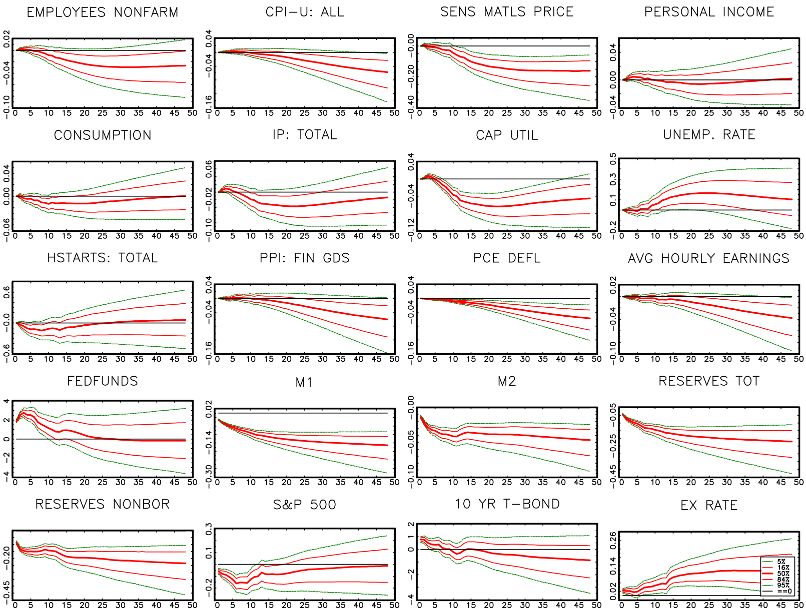
- We simulate the distribution of impulse responses using 16000 draws and plot median, 5th, 16th, 84th, and 95th quantiles

Shocks to factors




EMPLOYEES NONFARM
CPI−U: ALL
SENS MATLS PRICE
PERSONAL INCOME
CONSUMPTION
IP: TOTAL
CAP UTIL
UNEMP. RATE
HSTARTS: TOTAL
PPI: FIN GDS
PCE DEFL
AVG HOURLY EARNINGS
FEDFUNDS
M1
M2
RESERVES TOT
RESERVES NONBOR
S&P 500
10 YR T−BOND
EX RATE
5%
16%
50%
84%
95%
==0


EMPLOYEES NONFARM
CPI−U: ALL
SENS MATLS PRICE
PERSONAL INCOME
CONSUMPTION
IP: TOTAL
CAP UTIL
UNEMP. RATE
HSTARTS: TOTAL
PPI: FIN GDS
PCE DEFL
AVG HOURLY EARNINGS
FEDFUNDS
M1
M2
RESERVES TOT
RESERVES NONBOR
S&P 500
10 YR T−BOND
EX RATE
5%
16%
50%
84%
95%
==0


EMPLOYEES NONFARM
CPI–U: ALL
SENS MATLS PRICE
PERSONAL INCOME
CONSUMPTION
IP: TOTAL
CAP UTIL
UNEMP. RATE
HSTARTS: TOTAL
PPI: FIN GDS
PCE DEFL
AVG HOURLY EARNINGS
FEDFUNDS
M1
M2
RESERVES TOT
RESERVES NONBOR
S&P 500
10 YR T–BOND
EX RATE
5%
16%
50%
84%
95%
==0

Monetary policy shock

- We analyze the effects of a shock to the federal funds rate using the "medium" dataset of Banbura, Giannone, Reichlin (2010)
- We set the system rank to 3 and the lag length to 13.
- We split the data in fast and slow moving variables, and use standard Cholesky identification scheme

Data

Variable	Short name	Type
Employees on nonfarm payroll	EMP: TOTAL	slow
CPI, all items	CPI-U: ALL	slow
Index of sensitive material prices	SENS MAT'LS PRICE	slow
Personal income	PI	slow
Real Consumption	CONSUMPTION	slow
Industrial Production Index	IP: TOTAL	slow
Capacity Utilization	CAP UTIL	slow
Unemployment rate	U: ALL	slow
Housing starts	HSTARTS: TOTAL	slow
Producer Price Index (finished goods)	PPI: FIN GDS	slow
Implicit price deflator for personal consumption expenditures	PCE DEFL	slow
Average hourly earnings	AHE: GOODS	slow
Federal Funds, effective	FEDFUNDS	fast
M1 money stock	M1	fast
M2 money stock	M2	fast
Total reserves of depository institutions	RESERVES TOT	fast
Nonborrowed reserves of depository institutions	RESERVES NONBOR	fast
S&P's common stock price index	S&P 500	fast
Interest rate n treasury bills, 10 year constant maturity	10 YR T-BOND	fast
Effective Echange rate	EX RATE: AVG	fast

Responses to monetary policy shock

- The impulse responses are based on the representation:

$$Y_t = \{A(L)[I - B_0 A(L)]^{-1} B_0 + I\} \Lambda^{-1} \epsilon_t^* \tag{45}$$

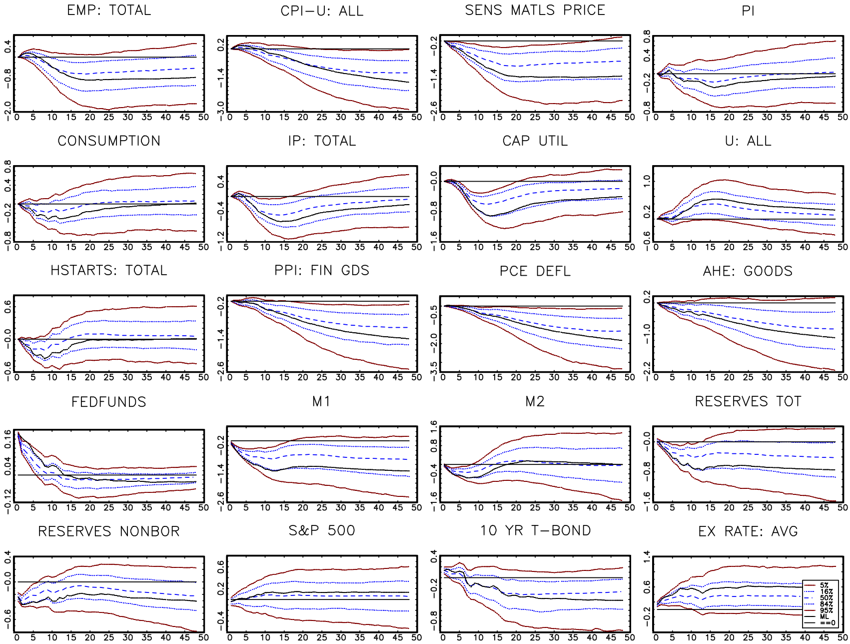
where ϵ_t^* are structural shocks and Λ^{-1} is the Cholesky factor of the reduced form shocks ϵ_t

- The resulting s-period ahead response is:

$$\Psi_s = A_1 B_0 \Psi_{s-1} + ... + A_{\min(s,p)} B_0 \Psi_{s-\min(s,p)};\ s > 0 \tag{46}$$

with $\Psi_0 = \{A(0)[I - B_0 A(0)]^{-1} B_0 + I\} \Lambda^{-1} = \Lambda^{-1}$.

- We simulate the distribution of impulse responses using 5000 draws and plot median, 5th, 16th, 84th, and 95th quantiles
- We also compute the impulse response using ML. Bootstrapping yielded very large bands.


EMP: TOTAL
CPI−U: ALL
SENS MATLS PRICE
PI
CONSUMPTION
IP: TOTAL
CAP UTIL
U: ALL
HSTARTS: TOTAL
PPI: FIN GDS
PCE DEFL
AHE: GOODS
FEDFUNDS
M1
M2
RESERVES TOT
RESERVES NONBOR
S&P 500
10 YR T−BOND
EX RATE: AVG
5%
16%
50%
84%
95%
ML
==0

Conclusions

- We have shown how to use RR-VARs for structural analysis
- We have discussed classical and Bayesian estimation and rank determination
- We have provided two empirical applications, focusing on VAR and factor style identification approaches
- Overall the method looks general, simple, and well performing, so promising for empirical analyses with large datasets

Geweke harmonic mean estimator

- The marginal data density $p_r(Y) = \int p(Y|\theta)p(\theta)d\theta$ is approximated numerically by using Geweke's (1999) modified harmonic mean estimator. In particular, by collecting all the coefficients in the vector $\theta = (A, \Sigma, \widetilde{B}_0)$ and considering the simulated posterior $\{\theta_j, \}_{j=1}^{m} = \left\{A_j, \Sigma_j, \widetilde{B}_0\right\}_{j=1}^{m}$ the estimator is:

$$\hat{p}(Y) = \left[\frac{1}{d}\sum_{m=1}^{d}\frac{f(\theta^m)}{p(Y|\theta^m)p(\theta^m)}\right]^{-1},$$

  where $f(\cdot)$ is a truncated multivariate normal distribution calibrated using the moments of the simulated posterior draws $\{\theta\}_{m=1}^{d}$. See Geweke (1999) for details.