Structural FECM: Cointegration in large-scale structural FAVAR models

Anindya Banerjee
(University of Birmingham)

Massimiliano Marcellino
(Bocconi University and CEPR)

Igor Masten
(University of Ljubljana)

Institut za javne financije
18. listopad, 2013

Motivation

- *Cointegration in factor-augmented VAR models*:
  - FECM - Factor-augmented error-correction model - Banerjee and Marcellino (2009)
  - FECM outperforms the FAVAR and the standard error-correction model in forecasting - (Banerjee, Marcellino and Masten, 2013)
- This paper: Structural FECM - implications of cointegration in large systems of non-stationary variables
- *Structural modelling in large systems*
  - Bernanke, Boivin and Eliasz (2005) introduce the **FAVAR**
  - Stock and Watson (2005)
  - Forni, Giannone, Lippi and Reichlin (2009) - fundamentalness of structural representations
  - These and other similar applications work with I(1) data transformed to I(0)
  - Neglected potential cointegration between factors and variables

Contributions of this paper

- Derivation of the FECM from the DFM representation of non-stationary data
- Derivation of the moving-average representation of the FECM - extension of the Granger representation theorem
- Structural FECM: first discussion of long-run identification schemes for non-stationary DFMs
  - Forni et al. (2009), Stock and Watson (2005) - long-run identification in I(0) panel
  - Eickmeier (2009) - I(1) panel and sign restrictions
- Analysis of the importance of the error-correction mechanism in large panels through empirical examples and simulation experiments

Structure of presentation

- Basic idea of the FAVAR
- Representation of the Factor-augmented Error-Correction Model
- Structural analysis:
  - Identification with long-run restrictions
  - Identification with short-run/contemporaneous restrictions (omitted)
- Empirical applications
- Simulation experiments
- Conclusions

FAVAR

Bernanke, Boivin and Eliasz, 2005

- Consider that the economy is driven by $\mathbf{Y_t}$ and $\mathbf{F_t}$
- $\mathbf{Y_t}$ observed, while $\mathbf{F_t}$ unobserved by econometrician
- $\mathbf{Y_t}$ and $\mathbf{F_t}$ follow a VAR
- Many observed I(0) indicator variables $\mathbf{X_t}$
- Assume factor structure

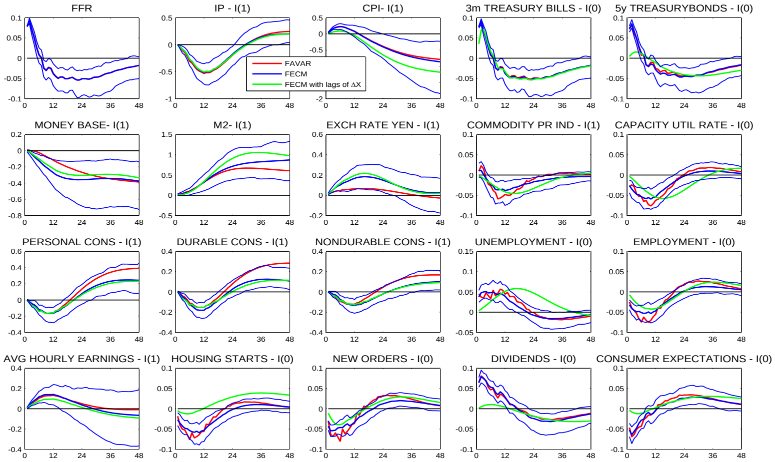
$$X_t = \Lambda^f F_t + \Lambda^y Y_t + e_t$$

FAVAR

BBE estimation

- Two approaches to estimation: likelihood-based Gibbs sampling or two-step approach with principal component analysis
- Two-step approach:
  - Estimate the space spanned by $\mathbf{Y_t}$ and $\mathbf{F_t}$ by PCA
  - Impose one of the factors to be observed and equal to the FFR, rotate the rest such that it is orthogonal to FFR, denote by $\tilde{F}_t$
  - Estimate the VAR for $[\tilde{F}_t', Y_t]'$, where FFR ordered last.
  - Estimate impulse responses of the FFR and factors
  - Regress each X_{it} on FFR and rotated factors, plug in the IRs of factors and get IRs of X_{it}

FAVAR - What do we get?



Dynamic factor model for non-stationary data

$$\begin{aligned} X_{it} &= \sum_{j=0}^{p} \lambda_{ij} F_{t-j} + \sum_{j=0}^{m} \varphi_{il} c_{t-l} + \varepsilon_{it} \\ &= \lambda_i(L) F_t + \varphi_i(L) c_t + \varepsilon_{it}, \end{aligned} \tag{1}$$

- Observables: $X_{it}, i = 1, \ldots, N, i = 1, \ldots, T$.
- T, N large.
- **Common trends** F_t - r_1-dimensional vector of I(1) factors,
- **Common cycles** c_t - r_2-dimensional vector of I(0) factors,

Dynamic factor model for non-stationary data

- $\lambda_i(L)$ - lag polynomial of order p,
- $\varphi_i(L)$ - lag polynomial of order m,
- $E(\lambda_i \varepsilon_{is}) = E(\varphi_i \varepsilon_{is}) = 0$ for all t, i and s.
- Strict DFM assumption:
  - ε_{it} is allowed to be serially correlated: $\varepsilon_{it} = \gamma_i(L)\varepsilon_{it-1} + v_{it}$,
  - $E(\varepsilon_{it}, \varepsilon_{js}) = 0$ for all i, j, t and s, $i \neq j$.
  - Stock and Watson (2005): empirically rejected, but quantitatively of limited importance.

DFM in static form

- Note that

$$\begin{aligned} \lambda_i(L) F_t &= \lambda_{i0} F_t + \lambda_{i1} F_{t-1} + \cdots + \lambda_{ip} F_{t-p} \\ &= \tilde{\lambda}_{i0} F_t - \tilde{\lambda}_{i1} \Delta F_t - \cdots - \tilde{\lambda}_{ip} \Delta F_{t-p+1} \end{aligned}$$

where

$$\tilde{\lambda}_{ik} = \lambda_{ik} + \lambda_{ik+1} + ... + \lambda_{ip}, \;\; k = 0, ..., p$$

- Define also

$$\tilde{\Phi}_i = [\phi_{i0}, ..., \phi_{im}]'$$

and

$$\begin{aligned} \Lambda_i &= \tilde{\lambda}_{i0} \\ \Phi_i &= \left[-\tilde{\lambda}_{i1}, \ldots, -\tilde{\lambda}_{ip}, \tilde{\Phi}_i\right] \\ G_t &= \left[c_t', c_{t-1}', ..., c_{t-m}', \Delta F_t', ..., \Delta F_{t-p}'\right]' \end{aligned}$$

DFM in static form

- DFM in static form

$$X_t = \Lambda_0 F_t + \Phi G_t + \varepsilon_t \quad (2)$$

- ε_t serially correlated

DFM in ECM form

- If we remove serial correlation from ε_t we get

$$X_t = \Gamma(L)\Lambda_0 F_t + \Gamma(L)\Phi G_t + \Gamma(L) X_{t-1} + v_t \qquad (3)$$

- With convenient factorization

$$\Gamma(L) = \Gamma(1) - \Gamma_1(L)(1-L),$$

- ...and some manipulation we can obtain the **DFM in ECM form**

$$\Delta X_t = \underbrace{-(I-\Gamma(1))(X_{t-1} - \Lambda F_{t-1})}_{\text{Omitted in the FAVAR}} + \Lambda\Delta F_t + \Gamma_1(L)\Lambda\Delta F_{t-1}$$
$$\Phi G_t - \Gamma(1)\Phi G_{t-1} + \Gamma_1(L)\Phi\Delta G_{t-1} - \Gamma_1(L)\Delta X_{t-1} + v_t$$

The VAR process for factors

- F_t are random walks

$$F_t = F_{t-1} + \varepsilon_t^F$$

- c_t are stationary common cycles $|\rho| < 1$

$$c_t = \rho c_{t-1} + \varepsilon_t^c$$

- ε_t^F and ε_t^c are correlated invertible moving-average processes
- By inverting the moving-average processes for factor innovations, we get a conventional VAR for the factors

$$\begin{bmatrix} F_t \\ G_t \end{bmatrix} = \begin{bmatrix} M_{11}(L) & M_{12}(L) \\ M_{21}(L) & M_{22}(L) \end{bmatrix} \begin{bmatrix} F_{t-1} \\ G_{t-1} \end{bmatrix} + Q \begin{bmatrix} u_t \\ w_t \end{bmatrix}$$

- Q accounts for dynamic singularity of G_t

Estimation

- Our DFM uses the same (stricter) set of assumptions as Bai (2004)
- Number of $I(1)$ and total number of factors determined by criteria developed in Bai (2004) and Bai and Ng (2002)
- Space spanned by the I(1) and I(0) factors can be consistently estimated by principal components from the I(1) data in levels
- Given strict DFM assumption remaining parameters estimated equation by equation.
- No generated-regressors problem (Bai, 2004) provided $N >> T$
- With a simulation experiment we demonstrate that the method successfully retrieves the factor space and impulse responses even in small samples ($T, N < 50$)

Non-stationary data and the FAVAR

- FAVARs in Bernanke et al. (2005), Stock and Watson (2005) and Forni et al. (2009) are estimated on data differenced to I(0)
- Omission of the error-correction term:
  - Non-invertible MA component:

$$\Delta X_t = \Lambda_0 \Delta F_t + \Phi \Delta G_t + \Gamma(L) \Delta X_{t-1} + \Delta v_t \tag{4}$$

  - Note that the error-correction term has a factor structure:
    - Omitted EC term can be proxied by inclusion of lags of I(0) factors - theoretically infinitely many
    - Problematic in applied work
- Empirical relevance: BBE dataset contains 120 series. 77 non-stationary. Loading to the EC term statistically significantly different from zero in 53 cases.

Moving average representation of the FECM

Granger representation theorem for the FECM

- Rewrite the factors VAR as

$$\begin{bmatrix} \Delta F_t \\ \Delta G_t \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_M \end{bmatrix} \begin{bmatrix} 0 & I_{r_2} \end{bmatrix} \begin{bmatrix} F_{t-1} \\ G_{t-1} \end{bmatrix} + \begin{bmatrix} M^*_{11}(L) & M^*_{12}(L) \\ M^*_{21}(L) & M^*_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta F_{t-1} \\ \Delta G_{t-1} \end{bmatrix} + Q \begin{bmatrix} u_t \\ w_t \end{bmatrix}$$

- Combine with the error-correction representation of the DFM

$$\Delta X_t = \tilde{\alpha} \left(X_{t-1} - \Lambda F_{t-1} - \Phi G_{t-1} \right) + \Lambda \Delta F_t + \Phi \Delta G_t + \Gamma_1 (L) \left(\Lambda \Delta F_{t-1} + \Phi \Delta G_{t-1} \right) - \Gamma_1 (L) \Delta X_{t-1} + v_t, \quad (5)$$

Moving average representation of the FECM

Granger representation theorem for the FECM

Then we can derive the GRT for the FECM as

$$\begin{bmatrix} X_t \\ F_t \\ G_t \end{bmatrix} = \underbrace{\begin{bmatrix} \Lambda \\ I_{r_1} \\ 0_{r_2 \times r_1} \end{bmatrix} \omega \sum_{i=1}^{t} u_t}_{\text{Common trends}} + \underbrace{C_1(L) \begin{bmatrix} v_t + [\Lambda, \Psi] Q [u_t', w_t']' \\ Q \begin{bmatrix} u_t \\ w_t \end{bmatrix} \end{bmatrix}}_{\text{Stationary part}}$$

where

$$\omega = \left[(I_{r_1} - M_{11}^{*}(1)) \right]^{-1}$$

has full rank

Identification of structural shocks based on long-run restrictions

- As is standard in SVAR analysis we assume that structural dynamic factor innovations are linearly related to reduced form innovations.

$$\begin{bmatrix} \eta_t \\ \mu_t \end{bmatrix} = H \begin{bmatrix} u_t \\ w_t \end{bmatrix}$$

- H is full-rank matrix
- η_t - r_1 permanent structural dynamic factor innovations
- μ_t - r_2 transitory structural dynamic factor innovations.

Identification of structural shocks based on long-run restrictions

- From the MA representation we see that the permanent effects are

$$\Lambda \omega u_t$$

- Assume that the **permanent effects of structural shocks** are

$$\Lambda^* \omega^* \eta_t$$

- Consider identifying real and nominal shocks as in Blanchard and Quah (1990).
- Identifying restrictions: nominal shocks have no long-run effect on real variables
- King, Plosser, Stock and Watson (1991) apply this logic to the cointegrated VAR
- We extend to large N

Identification of structural shocks based on long-run restrictions

- Partition X_t such that N_1 real variables ordered first, remaining $N_2 = N - N_1$ are ordered last.
- Assume $\Lambda^* \omega^*$ lower block diagonal, which implies ...
- ... both Λ^* and ω^* lower block diagonal

Identification of structural shocks based on long-run restrictions

Estimation of Λ^*

- The restricted loading matrix Λ^*:

$$\Lambda^* = \begin{bmatrix} \Lambda^*_{11} & \mathbf{0} \\ \Lambda^*_{21} & \Lambda^*_{22} \end{bmatrix}$$

- Λ^*_{11} and Λ^*_{21} estimated as loadings to the first factor - F^r_t - extracted from the whole dataset
- Estimate the residuals ε^r_t from a projection of X_t on F^r_t.
- Λ^*_{22} estimated as loadings to the $(r_1 - 1)$ factors F^n_t - extracted from the lower N_2-dimensional block of ε^r_t.

Identification of structural shocks based on long-run restrictions

Estimation of ω^*

- From the factors VAR

$$\hat{\omega} = \left[(I_{r_1} - M^*(11))\right]^{-1}.$$

- Estimate ω^* from the long-run covariance matrix

$$\omega E(u_t^F u_t^{F\prime})\omega' = \omega^* E(\eta_t \eta_t')\omega^{*\prime} = \omega^* \omega^{*\prime} \qquad (6)$$

where $\eta_t = [\eta_t^{r\prime}, \eta_t^{n\prime}]'$ are the structural innovations and ω^* is lower block diagonal.

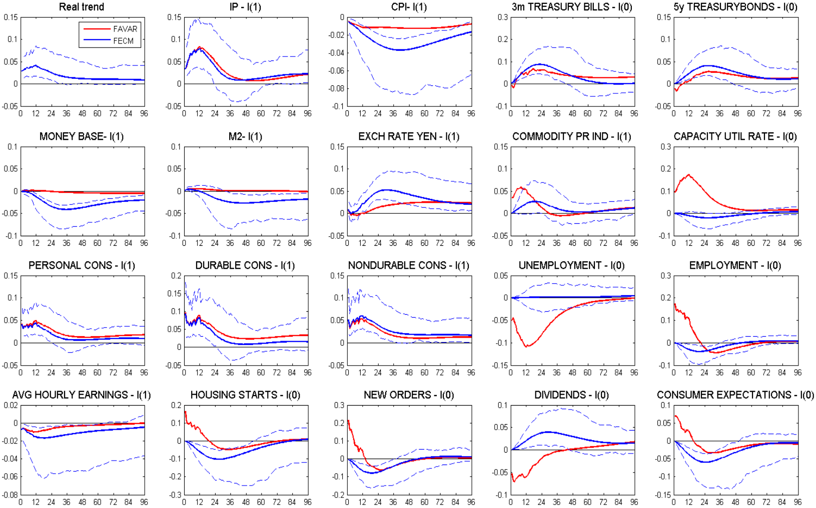
Empirical illustration

Data

- BBE dataset: 120 monthly US series for 1959 - 2003
- Extract 4 factors from data in levels (determined by information criteria of Bai (2004) and Bai and Ng (2002)
- Apply identification scheme from above
- Boostrapped 90% confidence intervals

Empirical example

Impulse responses to permanent real shock



Effect of ommiting the ECM term in the FAVAR

Percentage of FAVAR responses outside the FECM confidence intervals

Variables		Horizon					
	CI	12	24	36	48	60	72
All	67	32.5	55.8	63.6	57.1	48.1	41.6
	90	14.3	35.1	40.3	35.1	33.8	26.0
Output	67	5.6	22.2	33.3	27.8	22.2	16.7
	90	0.0	5.6	5.6	5.6	5.6	5.6
Employment	67	29.4	58.8	70.6	58.8	35.3	23.5
	90	0.0	29.4	41.2	17.6	17.6	11.8
Consumption	67	0.0	20.0	20.0	20.0	20.0	20.0
	90	0.0	0.0	0.0	0.0	0.0	0.0
Orders	67	0.0	100.0	100.0	100.0	50.0	0.0
	90	0.0	0.0	0.0	0.0	0.0	0.0
Exchange rates	67	50.0	75.0	75.0	50.0	25.0	25.0
	90	25.0	50.0	50.0	25.0	25.0	25.0
Money	67	55.6	77.8	77.8	88.9	88.9	88.9
	90	33.3	66.7	66.7	77.8	77.8	77.8
Prices	67	50.0	100.0	100.0	100.0	100.0	100.0
	90	50.0	50.0	50.0	50.0	50.0	50.0
Wages	67	60.0	100.0	100.0	93.3	93.3	93.3
	90	26.7	80.0	93.3	93.3	86.7	53.3

Monte Carlo experiment

- Simulation experiment to analyse the effect of omitting the ECM term in the FAVAR
- Data generating process: FECM estimated on 77 I(1) variables from BBE dataset
- Effect of:
  - Strength of error-correction: $\alpha = [-0.25, -0.50, -0.75]$
  - T dimension: $T = [250, 500]$
  - N dimension: $N = [50, 100]$

Monte Carlo results

Percentage of FAVAR responses outside the FECM confidence intervals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α	-0.50		-0.25		-0.75		-0.50		-0.50	
T	500		500		500		250		500	
N	100		100		100		100		50	
	Confidence interval coverage (%)									
Horizon	67	90	67	90	67	90	67	90	67	90
6	44.6	24.3	30.68	16.6	48.7	29.2	34.4	17.9	21.8	12.8
12	47.0	26.4	38.06	21.7	47.3	29.0	42.8	24.7	23.5	14.3
18	48.8	29.3	43.42	25.1	49.8	32.3	43.7	25.1	24.9	14.8
24	51.3	31.3	45.99	26.9	49.5	30.2	43.0	24.4	25.4	14.9
36	47.6	28.2	45.93	25.7	46.7	26.4	40.5	22.2	21.7	11.1
48	43.5	23.9	40.8	21.0	41.5	21.7	39.0	21.0	18.5	8.6
any	85.9	61.6	76.13	51.9	86.2	66.1	79.7	57.2	41.2	28.7
	Average % of periods IRs outside confidence interval									
	50.4	36.4	47.9	35.6	49.1	33.8	44.9	31.3	47.9	35.0

Conclusion

- This paper introduces the structural FECM
- We provide a generalization of the Granger representation theorem to large panels, which facilitates the analysis of structural shocks
- First application of the long-run identification scheme to cointegrated large-scale models
- Empirical and simulation evidence show that it is important to account for the error-correction mechanism in the FAVAR
- Responses to permanent real shocks compatible with DSGE evidence

Monetary policy shocks - FAVAR vs FECM