Roma June 2008. Banca d'Italia

DSGE Models in the Policy Environment

Marco Lippi. Comment on Jean Boivin and Marc Giannoni, Optimal Monetary Policy in a Data-Rich Environment.

Private sector:

$$\begin{bmatrix} Z_{t+1} \\ \\ \tilde{E}E_t z_{t+1} \end{bmatrix} = A \begin{bmatrix} Z_t \\ \\ z_t \end{bmatrix} + Bi_t + \begin{bmatrix} u_{t+1} \\ \\ 0 \end{bmatrix}$$

Z_t are predetermined variables.

z_t are forward-looking variables.

$I_t^p = \{Z_s,\ z_s,\ i_s,\ u_s,\ s \geq t;\ \Theta\}$ private sector information set.

$X_t^{cb} = \Lambda^{cb} \begin{bmatrix} Z_t \\ z_t \end{bmatrix} + v_t$ indicators observed by the central bank; v_t is an error-in-variables vector.

$I_t^{cb} = \{X_s^{cb},\ i_s,\ s \geq t;\ \Theta\}$ central bank information set.

Equilibrium under optimal policy but imperfect information:

$$\begin{bmatrix} z_t \\ \bar{z}_t \end{bmatrix} = DS_t$$

$$S_t = GS_{t-1} + H\epsilon_t$$

where

$$\bar{z}_t = \begin{bmatrix} z_t \\ z_{t|t} \end{bmatrix}, \quad S_t = \begin{bmatrix} Z_t \\ \Xi_{t-1} \\ Z_{t|t} \\ \Xi_{t-1|t} \end{bmatrix}, \quad \epsilon_t = \begin{bmatrix} u_t \\ 0 \\ v_t \end{bmatrix}$$

Different state vectors S_t and equilibrium equations are obtained under different assumptions on the central bank behavior.

Econometrician: Estimation of parameters and states.

$$S_t \; = GS_{t-1} + H\epsilon_t$$

$$X_t \; = \Lambda S_t + e_t$$

where

$$X_t = \begin{bmatrix} X_{F,t} \\ X_{S,t} \end{bmatrix}, \quad e_t = \begin{bmatrix} e_{F,t} \\ e_{S,t} \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_F \Phi \\ \Lambda_S \end{bmatrix}$$

where $X_{F,t}$ are indicators of some of the variables in the model, whereas $X_{S,t}$ is a potentially large vector of indicators of variables belonging to the economy. This is where the models gets linked to large-dimensional dynamic factor models.

The authors insist on a very important feature of the model: unlike finite dimensional unobserved components models, if the number of variable tends to infinity, the unobserved components can be consistently estimated. I think that this point might be made more forcefully. Using the authors' exemple, assume that

$$f_t = \rho f_{t-1} + \eta_t$$

but

$$x_t = f_t + e_t = \frac{1}{1 - \rho L}\eta_t + e_t$$

is observed. Then

$$(1 - \rho L)x_t = \eta_t + (1 - \rho L)e_t = (1 - \tau L)\zeta_t.$$

$$(1-\rho L)x_t = \eta_t + (1-\rho L)e_t = (1-\tau L)\zeta_t.$$

The deep parameters ρ, σ_η^2 and σ_e^2 can be consistently estimated, under the condition that $\rho \neq 0$.

However, you can never consistently estimate f_t. That is, the space spanned by x_t, for $t = -\infty, \infty$ does not contain f_t, so that

$$f_t - \mathrm{Proj}(f_t \mid x_\tau,\ \tau = -\infty, \infty)$$

does not vanish.

But, suppose that instead of just x_t, many variables are available and

$$\begin{aligned} x_{1t} &= f_t + e_{1t} \\ x_{2t} &= f_t + e_{2t} \\ &\vdots \\ x_{nt} &= f_t + e_{nt} \end{aligned}$$

where the errors e_{jt} are orthogonal to one another. Then

$$\frac{1}{n}\sum_{j=1}^{n} x_{jt} = f_t + \frac{1}{n}\sum_{j=1}^{n} e_{jt}$$

and

$$\operatorname{var}\left(\frac{1}{n}\sum_{j=1}^{n} e_{jt}\right) = \frac{1}{n^2}\sum_{j=1}^{n} \sigma_j^2 \leq \frac{1}{n}\max_j \sigma_j^2$$

Thus f_t belongs to the space spanned by the observables for $n \longrightarrow \infty$.

$$\frac{1}{n}\sum_{j=1}^{n} x_{jt} = f_t + \frac{1}{n}\sum_{j=1}^{n} e_{jt}$$

and

$$\operatorname{var}\left(\frac{1}{n}\sum_{j=1}^{n} e_{jt}\right) = \frac{1}{n^2}\sum_{j=1}^{n} \sigma_j^2 \leq \frac{1}{n}\max_j \sigma_j^2$$

Thus f_t belongs to the space spanned by the observables for $n \longrightarrow \infty$.

The authors insist on efficiency of the large dimensional model over the small-dimensional. I would rather insist on consistency.

Estimation.

$$S_t = GS_{t-1} + H\epsilon_t$$

$$\begin{aligned} X_t &= \Lambda S_t + e_t \text{ Static representation} \\ &= B(L)\begin{bmatrix} u_t \\ v_t \end{bmatrix} + e_t \text{ Dynamic representation} \end{aligned}$$

u_t are structural shocks.

v_t are errors in indicators used by the central bank.

We call S_t the static factors and $[u_t \; v_t]'$ the dynamic (primitive) factors. In principle, as $n \to \infty$, the estimator used to recover S_t should not matter. So we can compare results obtained with the estimator used in the paper and the principal components estimators employed in large-dimensional dynamic factor models.

In particular, in the empirical application of the paper I see 6 dynamic factors, see p.23, 3 of them being structural, the other central bank errors of measurement. Now, 6 is quite a big number as compared to what is found in the large-dimensional DFM's.

Hallin and Liška (2005) find one dynamic factor with Euro-area quarterly data.

Bai and Ng (2007) find 4 dynamic factors with US monthly data (instead of 2 factors, as advocated by Giannone, Reichlin and Sala, 2005).

This is not conclusive of course, but there seems to be convergence on a very small number of dynamic factors, no more than four with monthly data. The problem is interesting per se. Is the number of factors really small or some of them fail to be revealed by the relationship to the observable variables ? Consider a very simple example

$$x_{it} = a_{i1}u_{1t} + a_{i2}u_{2t} + e_{it}$$

so we have two static and dynamic factors.

1. If all the x's load the factors in the same way, i.e. with proportional coefficients, then only a linear combination of the factors can be recovered using the x's.

2. Suppose there are two different linear combinations of the factors. Yet, both linear combinations must occur an infinite number of times, otherwise we are again in case 1.

$$x_{it} = a_{i1}u_{1t} + a_{i2}u_{2t} + e_{it}$$

so we have two static and dynamic factors.

1. If all the x's load the factors in the same way, i.e. with proportional coefficients, then only a linear combination of the factors can be recovered using the x's.

2. Suppose there are two different linear combinations of the factors. Yet, both linear combinations must occur an infinite number of times, otherwise we are again in case 1.

3. Thus, existence of factors does not automatically means that we can recover them (that they are identified). It is required that the variables x respond with asymptotic heterogeneity.

The empirical exercise. Estimation of the parameters, of the state vector S_t. Welfare gains.

I am not sure I understand all the details of the estimation procedure. However:

a. The estimation of the state vector S_t should asymptotically depend only on the variables x (principal components estimate consistently S_t). What is the advantage of the procedure proposed here and a two-step estimation consisting in estimating the common components of the variables X_t, including $X_{F,t}$? As long as the model is identified this should be equivalent ??

b. Estimation is conducted under historical policy by the central bank, which is neither optimal nor full info. Then welfare gains under alternative policies are computed.

b. Estimation is conducted under historical policy by the central bank, which is neither optimal nor full info. Then welfare gains under alternative policies are computed.

All the results go in the right direction, with the welfare loss decreasing when full information or optimality are introduced in the central bank decision criterion.

However, it appears that when optimal policy is introduced, then the welfare gain of full information is not as impressive as it is under the historical rule (p.29). The authors suggest that the introduction of additional shocks to the model should address this problem. But adding more shocks makes even more necessary an autonomous analysis of the dynamic dimension of the dataset X_t.

Summing up:

This is an impressive big-orchestra paper, putting together lines of research that have been developed with different aims and tools.

Further clarification of the relationship between dynamic factor models and DSGE can provide new insights and ideas in both lines of research.