Report on :

« A replication of "A Quasi-Maximum Likelihood Approach for Large Approximate Dynamic Factor Model" (Review of Economics and Statistics, 2012) »

by Riccardo J. Lucchetti and Ioannis A. Venetis

This very short paper presents a replication and an extension of the Doz *et al* (2012) (DGR hereafter) Monte-Carlo experiment. The general setup is presented in section 2, the replications in section 3 and the extensions in section 4.

In section 3, the authors obtain results which are very similar to those obtained by DGR. They use an identical DGP, with identical values for the parameters, but they run 5000 replications whereas DGR have run only 500 replications so that the results they obtain, reinforce the initial ones.

In section 4, the authors extend the model of DGR and introduce lags in the measurement equation, which is now :

$$x_t = \Lambda_0 f_t + \Lambda_1 f_{t-1} + \cdots + \Lambda_s f_{t-s} + e_t \quad (1)$$

whereas the dynamics of f_t is unchanged and is a VAR(p).

As it is well known, equation (1) admits the static representation

$$x_t = \Lambda F_t + e_t \quad (2)$$

where $\Lambda = (\Lambda_0\ \Lambda_1 \dots\ \Lambda_s)$ and $F_t = (f'_t\ f'_{t-1} \dots\ f'_{t-s})'$.

The model Monte-Carlo simulations use $q = 2$ dynamic factors (*i.e.* f_t has dimension 2) which enter with one lag in the measurement equation ($s = 1$). The number of static factors (*i.e.* the dimension of F_t) is $q(s + 1)$.

The authors propose to compare 5 estimations methods, which are listed page 8. Some remarks must be done before going further :

– Method 3 should be applied using $q(s+1)$ principal components and not $q \cdot s$ since F_t has dimension $q(s + 1)$. This might be a typo but this error also appears in the second and fourth paragraphs of page 10.
– Method 4 is irrelevant : why should we extract 1 principal component when we know that the static factor has dimension $q(s + 1) = 4$?
– Method 5 is irrelevant too : why should we extract 1 principal component and take its first lag when we know that the dynamic factor has dimension $q = 2$ and should indeed appear with one lag ?

The authors are aware of the two last problems and say (page 10) that it is a way to see how the PC estimators behave under misspecification. But this section aims at comparing QML estimators with other estimators. Thus, if the dimension of the factors is well specified for the QML estimator, it should be also well-specified for any PC estimator which is considered. The behavior of any of those estimation methods under misspecification could also be considered, but it is another topic.

Methods 1 and 2 are indeed an extension of DGR who did not consider the case where the factors enter the measurement equation with lags. As intialization is always an issue, when using numerical optimization methods, I think the authors should give more details on the way they proceed. They quote 3 reference papers, but these papers don't use the same methodology, so that we don't know which is the methodology used here.

In particular, for method 2, it is unclear how $\hat{f}_t$ is estimated using the two-step method of Doz *et al* (2011). Indeed in the first step of in Doz *et al* (2011) method, the factors are estimated by PCA, but the model is a static one (no lag in the measurement equation), which is not the case here. In particular, in the present case, using a PCA to extract only q factors in the first step would certainly not be the right way to proceed. If it is what has been done, it could explain the disappointing results obtained for TS.

Overall, this paper presents not only a replication but also an extension of Doz *et al* (2012) simulations results. However in the extension part : the number of factors in methods 4 and 5 (and possibly in method 3 if $q \cdot s$ instead of $q(s+1)$ is not a typo) should be modified and the initialization procedures used in methods 1 and 2 should be made more precise.