Discussion of "A multiple testing approach to the regularisation of large sample correlation matrices"

Massimiliano Marcellino [1]

ECB Forecasting Conference, 18-19 June 2018

[1] Bocconi University, IGIER and CEPR

Summary

- Estimation of large covariance matrices important both in finance (portfolio construction, risk management, ...) and in macro (impulse response analysis, measures of uncertainty), more used in the former than in the latter, pity...
- Parameter dimensionality issue, various approaches in the statistical literature (e.g., Pourahmadi (2011)): banding, tapering, thresholding, shrinking.
- This paper belongs to the "thresholding" approach but, rather than using cross-validation to select the threshold, it uses (multiple) testing -> good idea!
- Shrinkage (towards I_N) can be used in a second step to guarantee a positive definite matrix

Summary

- Theoretically, the paper shows convergence of the MT-based estimator (under sparsity of the true matrix) both in spectral norm and in Frobenius norm, deriving the proper rates (which depend on N, T, and extent of dependence of the variables)
- In terms of finite sample performance, a set of Monte Carlo experiments find good performance, also wrt alternative methods (Bickel and Levina (2008, BL), Cai ad Liu (2011, CL), Ledoit and Wolf (2004, LW)).
- Overall a very nice paper with an interesting and relevant contribution!

Some comments on the MC

- Perhaps the choice of the design should be better motivated. In particular, is it close to empirical estimates of large correlation matrices in finance/macro?
- It would be interesting to also consider larger values of N (now max is $N = 200$) and larger $N/T > 1$ ratios (now max is $(N = 200)/(T = 100) = 2$), as estimation is particularly problematic when $N/T > 1$.
- From Tables 2 and 3, differences wrt to CL do not seem so large, and MC standard errors seem large enough to make them not statistically significant (test?), is it so? Perhaps try more simulations (100 now) only for this comparison?
- Differences wrt CL increase substantially when comparing "shrinked" versions of the estimators, is there an intuitive explanation?

Some comments on the empirical application

- Where is it???
- I am particularly interested in the application to de-factored observations (Fan et al. (2011, 2013)) or "de-VARed" observations (Carriero, Clark and Marcellino (2018, CCM-JoE)), as often in macro/finance we are interested in variance matrices of residuals (from large models).
- You mention MC results on de-factored observations are available upon request, are they good?
- Theoretical results? My guess is that if the model generating the residuals is correctly specified, your results remain valid (perhaps under slightly stricter conditions to ensure that use of residuals instead of errors does not matter). If the model is mis-specified, there can be problems. For example, if you use fewer factors than true, the omitted factor will create a non-sparse correlation matrix, same for omitted variables in VARs.
- CCM-JoE highlight another possible issues: even in large VAR models variances are time-varying...

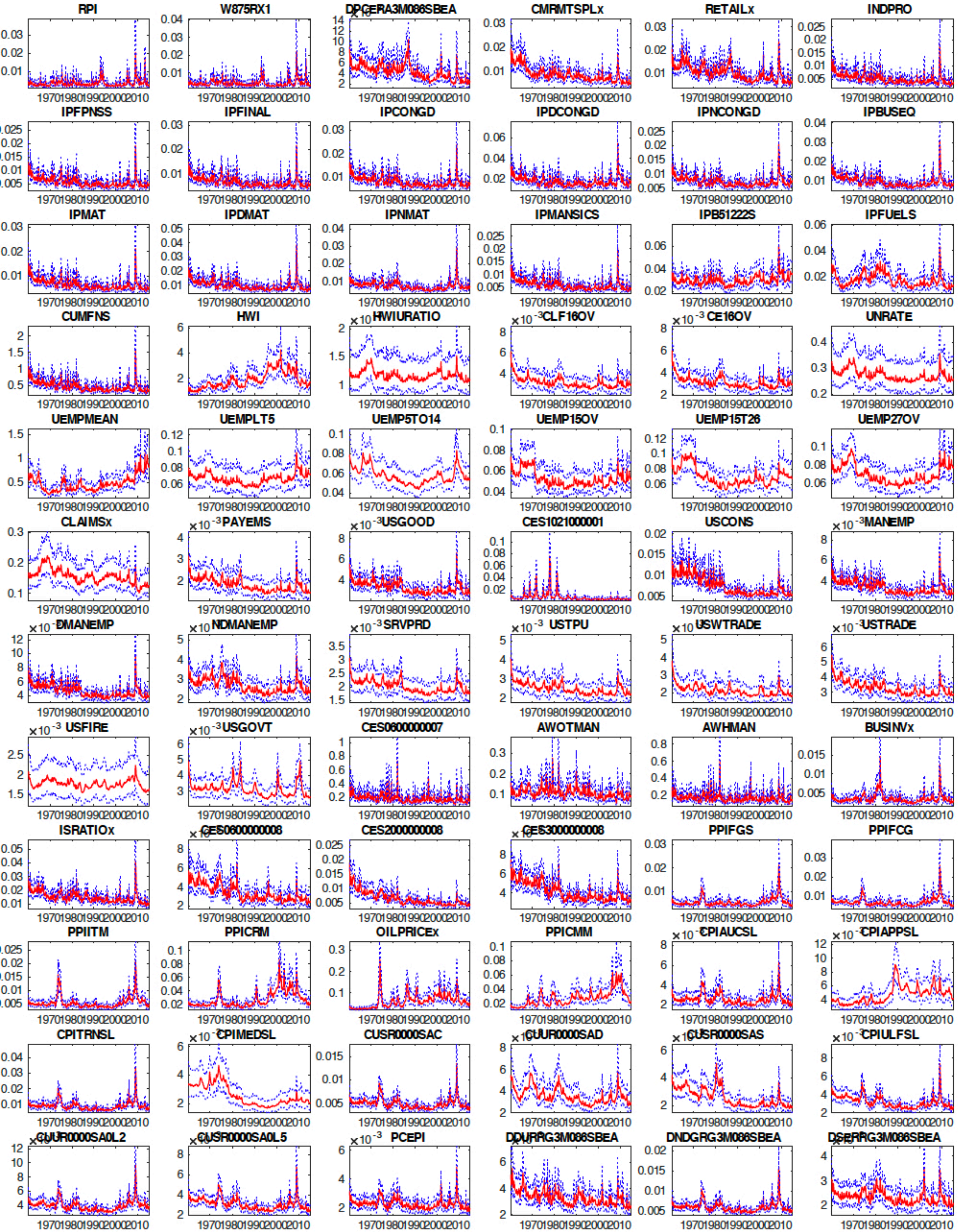

Figure 7: Posterior distribution of volatilities (diagonal elements of Σ_t), slow variables.

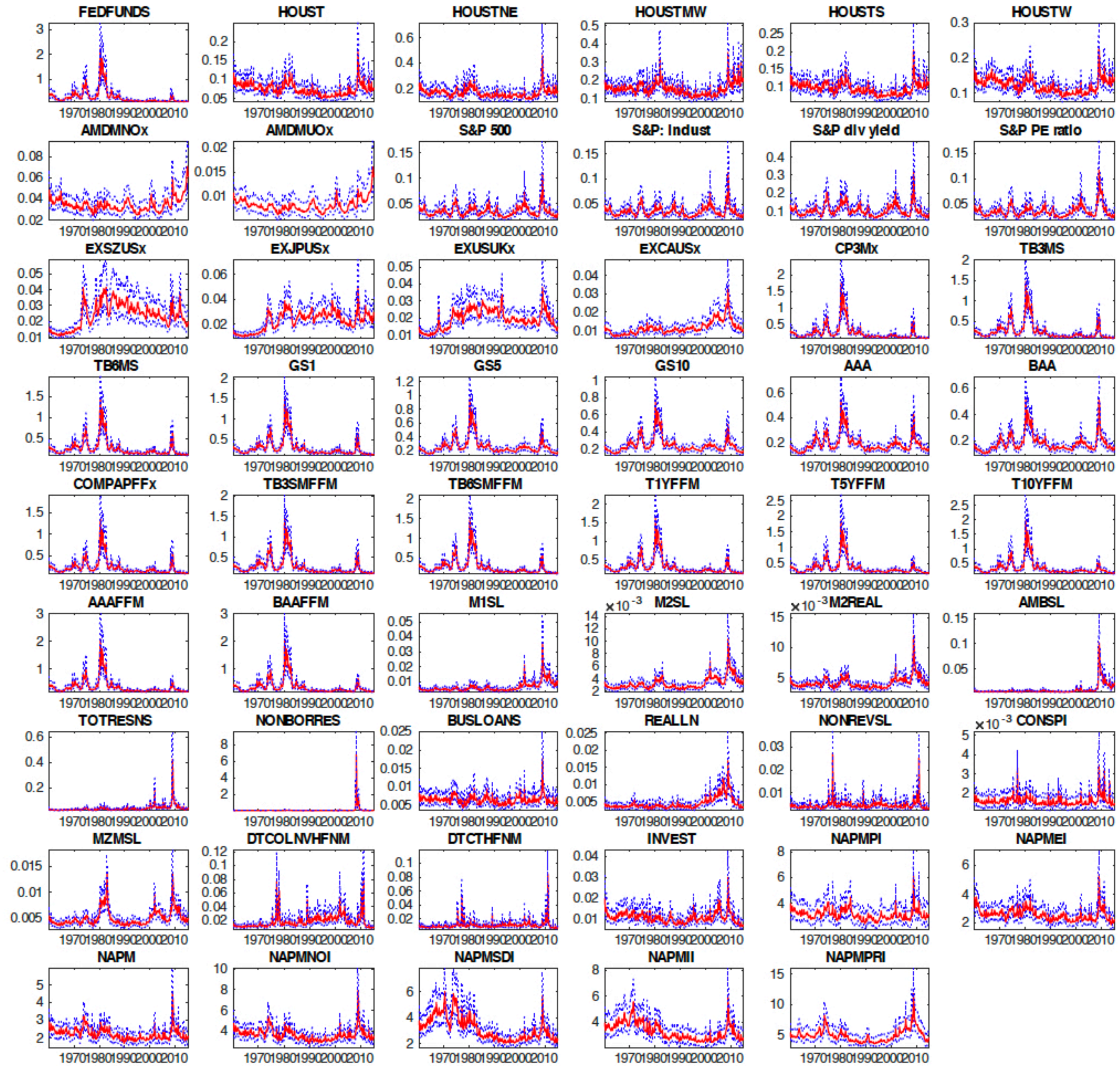

Figure 8: Posterior distribution of volatilities (diagonal elements of Σ_t), fast variables.

PCA of the variance matrix of the shocks to volatilities