The Making of "Estimation of Common Long–Memory Components in Cointegrated Systems"

JESÚS GONZALO
Department of Economics, U. Carlos III de Madrid

The eighties were very good years for music as well as econometrics. In time-series econometrics, the first half of that decade was dominated by research on unit roots while cointegration was the queen of the second half. Estimation and testing of a cointegrated system were the key questions to answer.

When I started my dissertation at the end of the eighties, under the supervision of Clive Granger and Robert Engle, you could sense that everyone was of the opinion that the testing problem of the cointegration rank had been solved by Johansen (1988). Johansen applied reduced rank regression techniques to the following error correction model (ECM) (Granger's old notation is used to keep the spirit of the eighties)

$$\Delta X_t = \gamma \alpha' X_{t-1} + \varepsilon_t \qquad (1)$$

and obtained the corresponding asymptotics. Nevertheless, estimation was still under close scrutiny, and new estimators appeared on a regular basis in the literature. It was also the case however that different methods could lead to very different estimates of the cointegrating vector, and this was the reason Clive and Rob wanted me to start working on the problem. In Gonzalo (1994), I obtained the asymptotic distribution of the principal components (method proposed in Stock and Watson 1988) and canonical correlations (between the vector X_t and X_{t-1} proposed by Bossaerts 1988) based estimators of cointegrating vectors and compared their performance to the most popular alternatives at the time: ordinary least square (Engle and Granger 1987), nonlinear least square (Stock 1987), and maximum likelihood (Johansen 1988). The latter was the clear winner once the dispersion of the finite sample distributions was measured by the interquartile range instead of the simple variance.

By that time, estimation and testing could be considered almost "dead topics" from a Ph.D. student's research perspective. However, the most important question about cointegration was still very much open. Namely *what makes a set of*

e-mail: jesus.gonzalo@uc3m.es

variables to be cointegrated? I felt the answer to this question could become the third chapter of my dissertation. We knew that cointegration occurs if and only if there is a common stochastic trend representation. Stock and Watson (1988) showed a possible way of obtaining the common trends. Granger, however, did not like much the fact that the common trends were taken to be random walks. He believed that trends could follow more general processes. As it is well known he was very good on linear combinations and he therefore had the idea that a cointegrating system could be characterized by only two sets of linear combinations: the cointegrating ones ($Z_t = \alpha' X_t$) and the ones driving the system in the long run. For Clive, a good candidate for the latter was $\gamma_\perp X_t$ because Z_t will not cause them in the long run.

Since the beginning of my dissertation, I was interested on cointegration as a dimensionality reduction technique, useful to analyze large data sets. This was something that Clive thought had not been fully exploited yet. Reading through the literature on factor models, it became clear that if we could make the common factors or common trends observable, we could gain valuable insight on the factors while at the same time helping the identification process. One possibility was to force them to be a linear combination of contemporaneous $X_t's$ but the question was *which* linear combination? The one Granger had in mind? Why not $\alpha'_\perp$ or any other outside the space spanned by α? These questions led me to specify a factor model imposing only the requirement that the $I(1)$ common factors be linear combinations of X_t, say

$$X_t = A_1 F_t + H_t \qquad (2)$$

with $F_t = B_1 X_t$ a common permanent $I(1)$ component and H_t a transitory $I(0)$ component. From Equation (2), $H_t = (I - A_1 B_1) X_t$. Since this component is assumed to be $I(0)$, it had to be formed by the cointegrating relationships: $H_t = A_2 \alpha' X_t$. I did not want to assume both components (permanent and transitory) to be orthogonal as it was done in the factor model literature because it was too strong of an assumption in the present context. After all, why shouldn't the growth rate of the trend affect the cycle component? It was clear however that the converse should not be possible in the long run if we wanted H_t to be a transitory component of X_t. It seemed therefore that causality instead of orthogonality was the right assumption. The latter satisfies the symmetry property while the former can work in only one direction. At this point, I perfectly understood what Granger had originally in mind: $F_t = B_1 X_t = \gamma'_\perp X_t$. Any other linear combination, like $F_t = \alpha'_\perp X_t$, would force H_t not to be transitory. As everybody can imagine, he was 300 steps ahead of me. I was very stubborn and wanted to get convinced by myself. Clive never mentioned that he was right after all and that $\gamma_\perp$ was the right linear combination! In all the weekly meetings I had with him, I tried to raise possible problems with his proposed linear combination but he never despaired. I think he knew perfectly that at least in the long run he was right. I am sure he left me to prove it as a part of my graduate learning process. The rest (existence of the permanent–transitory

decomposition, estimation of $\gamma_\perp$, testing hypotheses on it, etc.) was left for the graduate student to finish and can be found in Gonzalo and Granger (1995).

For the testing part, we were very lucky that on that time Soren Johansen was working on his Johansen (1991) while he was visiting University of California San Diego (UCSD). Didn't I say that the eighties were good? For a UCSD graduate student, the second half of that decade was indeed terrific. James Davidson, David Hendry, Soren Johansen, Katarina Juselius, James Mackinnon, and Jan Magnus were just a few of the visitors we enjoyed.

To work with Clive was a pleasure and a luxurious experience. He was a master of making things simple. He was very generous with his time and always willing to share his ideas with everybody except in his "siesta" time. I will always regret not to have collaborated more with him; but one of his subtle thoughts was that once you graduate you have to fly by yourself. And I tried to do that. THANKS CLIVE!!

REFERENCES

Bossaerts, P. 1988. Common Nonstationary Components in Stock Prices. *Journal of Economic Dynamics and Control* 12: 347–364.

Engle, R. F., and C. W. J. Granger. 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55: 251–276.

Gonzalo, J. 1994. Five Alternative Methods of Estimating the Long-Run Equilibrium Relationships. *Journal of Econometrics* 60: 203–233.

Gonzalo, J., and C. W. J. Granger. 1995. Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics* 13(1): 27–35.

Johansen, S. 1988. Statistical Analysis of Cointegrating Vectors. *Journal of Economic Dynamics and Control* 12: 231–254.

Johansen, S. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian VAR Models. *Econometrica* 59: 1551–1580.

Stock, J. H. 1987. Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica* 55: 1035–1056.

Stock, J. H., and M. W. Watson. 1988. Testing for Common Trends. *Journal of the American Statistical Association* 83: 1097–1107.