

Comments on
"Estimating Systems of Trending Variables"

by Soren Johansen



I am very glad to have the opportunity to make some comments on this excellent survey about the literature on cointegration and error correction systems.

The author summarizes in a very clarifying way most of the latest literature on estimation and testing in cointegrated systems. He clearly succeeds in explaining the relationship between his approach, based on VAR models, and the common trends approach, based on VMA models, of Stock and Watson(1990) when the variables are $I(1)$ and $I(2)$. Furthermore, he emphasizes the advantages and disadvantages of using the optimal inference procedures, proposed by Phillips(1990) and others, for cointegrated systems. His comments contain good advice for empirical researchers.

In order to try to contribute to the understanding of some of the limits of this paper I am going to concentrate on what are, from my viewpoint, its weakest elements, knowing that a survey will never fully satisfy everyone.

Although I feel sympathetic with the idea of *defining $I(d)$ in a way that is not model dependent*, I think it is a difficult task. In particular, Johansen's proposal is not free of some important caveats as illustrated with the following examples:

1) Let x_t be the following *heteroskedastic random walk*, $x_t = x_{t-1} + e_t$, where e_t is a series of independent random variables with mean 0 and time varying but bounded variance (heteroscedastic). Since Δx_t is nonstationary, from definition 2.2 the author will conclude that x_t is not $I(1)$ while in fact there is no explosive component (trend) in either the mean or the variance of the first difference. This problem affects also his definition 2.4 of cointegration since it is based on the concept of $I(d)$.

2) Johansen's *definition of trending variable is confusing* since for example $x_t = c + bt + e_t$ will not be trending, following definition 2.3, because it is not the sum of an $I(d)$ process and a polynomial trend.

The concept of trend has a long history in time series analysis, see for example Anderson(1958). Although there is no universal way of understanding what a trend is, in what follows I propose a definition of trend, based on Escribano(1987), that captures most of the notions of time trends used in applied work.

Definition D1. The univariate stochastic process $\{x_t\}$ has a smooth time trend of order $h(i,t)$ in the i -th moment if:

- a) $E(x_t^i)$ is finite for finite t ,
- b) $E(x_t^i) = h(i,t)$ where $h(i,t)$ is a smooth time dependent function of t , and if
- c) $h(i,t)$ is unbounded as $t \rightarrow \infty$.

From the empirical viewpoint, the most important cases of trends are those of the mean and the variance. In general, a deterministic trend generates a smooth time trend in the mean while a stochastic trend generates a smooth time trend in the variance. Alternative definitions of trends might be based on the concept of long memory, see for example Granger and Hallman(1991).

$I(d)$ processes have trends in variance and these trends should be removed by differencing the series d times. To select the order d of integration, $I(d)$, we should always subtract the mean, $h(1,t)$, first.

The idea behind the ARIMA(p,d,q) models, see Box and Jenkins(1976), is that we should differentiate as many times as necessary to find an autocorrelogram that declines "fast". If this is not the case the spectrum at zero frequency, $f(0)$, will be unbounded.

Those comments motivate the following definition of an $I(d)$ process.

Definition D2. The univariate stochastic process $\{x_t\}$ is weakly integrated of order d , $I(d)$, $d=0,1,2,3,\dots$, if:

- a) $\{x_t - h(1,t)\}$ has a smooth time trend in variance, and
- b) if after differencing at least d times, $\Delta^d\{x_t - h(1,t)\}$,
 - b1) has no smooth time trend in variance
 - b2) the spectrum at 0 frequency is bounded, $0 < f(0) < \infty$
 - b3) and it is stationary of second order

Under definition D2, $\{x_t\}$ has a Wold's representation and so it has an ARIMA(p,d,q) representation. Conditions b1, b2 and b3 on $\Delta^d\{x_t - h(1,t)\}$ are related to the conditions for short-range dependence, see Lo(1991) and Escribano(1987).

From the definition of $I(d)$, $d=0,1,2,\dots$, processes we conclude that the origin of the nonstationarity in the series is the existence of d unit roots. However, we can have $I(0)$ processes that are $I(0)$ but nonstationary. The typical case is a white noise process with bounded heteroskedasticity. The advantage of definition D2 is that it allows us to distinguish between different sources of nonstationarity. For example, if a series has a unit root and, after differencing once, satisfies conditions b1) and b2) but not b3), like the example of section 1 of this comment, then we will say that it is heterogeneous integrated of order 1, $HI(1)$.

3) Johansen's definition of $I(1)$ allows some components of the vector to be $I(0)$, as long as some of the others are at least $I(1)$. Hence the definition of cointegration might be empty of

content. The reason is twofold. First, it is always possible to find a cointegrating vector that has zero components corresponding to the $I(1)$ elements. Second, it is possible to include in the cointegrating vector elements that are stationary. To avoid these problems it will be enough to add to definition 2.3 the conditions that the cointegrating vector has zero components in all $I(0)$ variables and that the cointegrating vector has at least two elements different than zero.

4) Since Johansen's approach is based on VAR models, which are generally used to do impulse response analysis (Sims(1980), I am missing in this survey some comments on the impact of cointegration on impulse response analysis and on forecasting with VAR models.

Lütkepohl(1991, chapter 11), mentions that in some cases we cannot evaluate the long run effects of shocks through the total multipliers since they might not be finite. However, we can do impulse response analysis indicating, for example, that the effects of a shock may not die out even with long run horizons (permanent shocks).

Before finishing, I would like to congratulate Soren Johansen again for writing such a clarifying survey on cointegration. I am sure many people will benefit from reading it.

by Alvaro Escribano

Universidad Carlos III de Madrid

References

- Anderson T.W. (1958), The Statistical Analysis of Time Series, New York, Wiley and Sons.
- Box E.P and G.M. Jenkins (1976), Time series Analysis, Forecasting and Control, Holden Day, San Francisco.
- Escribano A. (1987), "Cointegration, Time Cotrend and Error Correction Systems: An Alternative Approach", CORE Discussion Paper 8715.
- Granger C.W.J. and J.J. Hallman (1991), "Long Memory Series with Attractors", Oxford Bulletin of Economics and Statistics, 53, 1, 11-26.
- Lo A. W. (1991), "Long-Term Memory in Stock Market Prices", Econometrica 59, 1279-1313.
- Lütkepohl H. (1991), Introduction to Multiple Time Series Analysis, Springer-Verlag.
- Sims C. A. (1980), "Macroeconomics and Reality", Econometrica 48, 1-48.