The Chang-Kim-Park Model of Cointegrated Density-Valued Time Series Cannot Accommodate a Stochastic Trend

Brendan K. Beare

In their article “Nonstationarity in Time Series of State Densities,” published in the Journal of Econometrics in 2016, Yoosoon Chang, Chang Sik Kim, and Joon Y. Park—referred to hereafter as CKP—propose a model of nonstationary cointegrated dynamics for a time series of probability density functions (hereafter “pdfs”). The main novelty of their approach is that at each point in time, the value taken by the time series in question is not a point in finite dimensional Euclidean space but rather a function, in fact a pdf, which we can view as an element of an infinite-dimensional vector space. While this necessarily complicates some of the mathematics, at an intuitive level their model can be understood to be an infinite-dimensional version of the usual model of cointegration. CKP fit their model to a time series of cross-sectional distributions of individual earnings, and a time series of intra-month distributions of stock returns.

Unfortunately, the dynamic model proposed by CKP has a rather serious flaw: Any time series of pdfs that satisfies their assumptions must necessarily be stationary, aside from a deterministic component. In particular, a stochastic trend driving long-run variation in their time series of pdfs—the defining feature of cointegration—cannot be present within their framework. The problem arises...
from the nonnegativity property of pdfs, a property that turns out to be incompatible with the postulated form of nonstationarity.

To build some intuition for what is going on, suppose we have two independent and identically distributed (“iid”) sequences of random variables, \((x_t, t \in \mathbb{N})\) and \((y_t, t \in \mathbb{N})\), which are independent of one another, with each \(x_t\) and \(y_t\) equal to 1 with probability \(\frac{1}{2}\) and equal to \(-1\) with probability \(\frac{1}{2}\). Construct a third sequence of random variables \((z_t, t \in \mathbb{N})\) by setting

\[ z_t = a + b \sum_{i=1}^{t} x_i + y_t \]

for each \(t \in \mathbb{N}\), for some constants \(a, b \in \mathbb{R}\). It should be clear that the time series \((z_t)\) is stationary (in fact iid) if \(b = 0\), and nonstationary if \(b \neq 0\). But what if I told you that each random variable \(z_t\) is guaranteed to be nonnegative? For \(z_t\) to be a nonnegative random variable we require that \(a \geq |b| t + 1\). This can only be true for all \(t \in \mathbb{N}\) if in fact \(b = 0\), and so we see that the requirement of nonnegativity forces \((z_t)\) to be stationary.

The situation in CKP’s paper is more complicated than in the simple example just provided, but it is the same tension between nonnegativity and nonstationarity that leads to the problem with their model. The problem is obscured by the fact that CKP work with demeaned pdfs. Though a demeaned pdf need not be nonnegative, it must nevertheless be bounded from below by a function that integrates to minus one. We will see that this leads to the same problem.

What follows is necessarily somewhat technical. I encourage readers who wish to learn more about time series analysis in infinite dimensional spaces to consult the monograph of Denis Bosq (2000).

Let \(K\) denote a compact interval of real numbers, and let \(B\) denote the Banach space of Lebesgue equivalence classes of integrable real valued functions on \(K\) equipped with the norm \(\|f\|_1 = \int_K |f|\). CKP’s time series is a sequence of pdf-valued random elements \((f_t, t \in \mathbb{N}_0)\) of \(B\), where \(\mathbb{N}_0 = \mathbb{N} \cup \{0\}\). For each \(t \in \mathbb{N}_0\) we let \(w_t = f_t - Ef_t\), with the expectation operator \(E\) defined as a Bochner integral in the usual way. We may informally view \(Ef_t\) as a pointwise Lebesgue integral \(E\mu_t(\cdot)\) taken for each \(\mu \in K\), though strictly speaking \(\mu_t(\cdot)\) is not well-defined for fixed \(\mu \in K\) as \(\mu_t\) is a Lebesgue equivalence class of functions.

Let \(H\) denote the Hilbert space of Lebesgue equivalence classes of square integrable real-valued functions on \(K\) that integrate to zero, with inner product \(\langle f, g \rangle = \int f g\) and corresponding norm \(\|\cdot\|_2\). CKP assume that \((w_t, t \in \mathbb{N}_0)\) is a sequence of random elements of \(H\). For \(t \in \mathbb{N}\) we let \(u_t = w_t - w_{t-1}\). CKP assume that there is an iid sequence \((\epsilon_t, t \in \mathbb{Z})\) in \(H\) with \(E\|\epsilon_0\|_2^2 < \infty\) and positive definite covariance operator \(\Sigma\) such that, for each \(t \in \mathbb{N}\), we have
\[ \mu_t = \sum_{k=0}^{\infty} \Phi_k(\epsilon_{t-k}). \]

Here \( (\Phi_k, k \in \mathbb{N}) \) is a sequence in \( \mathcal{L}(H) \), the space of bounded linear operators from \( H \) to \( H \). The sequence of operators is assumed to satisfy the summability condition \( \sum_{k=0}^{\infty} \|\Phi_k\|_{\mathcal{L}(H)} < \infty \), where \( \|\cdot\|_{\mathcal{L}(H)} \) is the operator norm. Under this summability condition, CKP observe that a version of the Beveridge-Nelson decomposition applies: We have

\[ w_t = w_0 - \tilde{\mu}_0 + \sum_{j=1}^{t} \Phi(\epsilon_j) + \tilde{\mu}_t \]  

(2)

for each \( t \in \mathbb{N} \), where \( \Phi = \sum_{k=0}^{\infty} \Phi_k \), and \( (\tilde{\mu}_t, t \in \mathbb{Z}) \) is the stationary sequence of random elements of \( H \) defined by \( \tilde{\mu}_t = \sum_{k=0}^{\infty} \tilde{\Phi}_k(\epsilon_{t-k}) \), with \( \tilde{\Phi}_k = - \sum_{j=k+1}^{\infty} \Phi_j \).

**Proposition:** Under the conditions just stated, \( \Phi = 0 \).

Before diving into the proof of this Proposition, let’s recall the very simple example used earlier to illustrate the tension between nonnegativity and nonstationarity. Compare equation (1) to the Beveridge-Nelson decomposition (2). We can think of \( z \) as playing the role of the demeaned pdf \( w \), of \( a \) as playing the role of the initial condition \( w_0 - \tilde{\mu}_0 \) of \( b \) as playing the role of the operator \( \Phi \), of \( x \) as playing the role of the innovation \( \epsilon \), and of \( y \) as playing the role of the stationary component \( \tilde{\mu} \). Just as a nonnegativity condition on \( z \) forced \( b \) to be equal to zero in our simple example, the proof of the above Proposition will demonstrate that the requirement that \( w \) be bounded from below by a function that integrates to minus one—specifically, by the negative of \( Ef_i \)—forces \( \Phi \) to be equal to zero.

**Proof of Proposition:** Each \( f_i \) is pdf-valued and hence nonnegative, so in view of the Beveridge-Nelson decomposition (2) we must have

\[ w_0 - \tilde{\mu}_0 + \sum_{j=1}^{t} \Phi(\epsilon_j) + \tilde{\mu}_t \geq -Ef_i \]

(3)

for all \( t \in \mathbb{N} \). Fix an arbitrary Borel set \( J \subseteq K \). A consequence of (3) is that we must have

\[ \frac{1}{\sqrt{t}} \int_J Ef_i + \frac{1}{\sqrt{t}} \int_J (w_0 - \tilde{\mu}_0) + \frac{1}{\sqrt{t}} \sum_{j=1}^{t} \Phi(\epsilon_j) + \frac{1}{\sqrt{t}} \tilde{\mu}_t \geq 0 \]

(4)

for all \( t \in \mathbb{N} \). Consider the behavior of each of the four terms on the left-hand side of the inequality (4) as \( t \to \infty \). The Bochner integral \( Ef_i \) is nonnegative and satisfies \( \int_E f_i = E\int f_i = 1 \), so we have \( 0 \leq \int_E f_i \leq 1 \) and hence \( t^{-1/2} \int_E f_i \to 0 \). Further, since \( \int_J (w_0 \to \infty) \)
\( - \hat{\mu}_0 \) and \( \int f \hat{\mu} \) are random variables whose law does not depend on \( t \)—the latter by stationarity—we have \( f^{-1} \int f (w_0 - \hat{\mu}_0) \to 0 \) and \( f^{-1} \int f \hat{\mu} \to 0 \) in probability. The central limit theorem for Hilbertian random elements (Bosq 2000, Theorem 2.7) ensures that \( f^{-1} \sum_{j=1}^t \Phi (\varepsilon_j) \) converges in law to \( Z \), a Gaussian random element of \( H \) with covariance operator \( \Phi \Sigma \Phi^* \), where \( \Phi^* \) is the adjoint to \( \Phi \). It then follows from the continuous mapping theorem that \( f^{-1} \sum_{j=1}^t \Phi (\varepsilon_j) \to f \int Z \) in law. Putting these results together with Slutsky’s theorem, we find that

\[
\frac{1}{\sqrt{t}} \int f \varepsilon_t + \frac{1}{\sqrt{t}} \int (w_0 - \hat{\mu}_0) + \frac{1}{\sqrt{t}} \sum_{j=1}^t \Phi (\varepsilon_j) + \frac{1}{\sqrt{t}} \int f \hat{\mu}_j \to f \int Z
\]

in law as \( t \to \infty \). Now, \( f \int Z \) is a centered Gaussian random variable, so the only way that (4) and (5) can simultaneously be true is if \( f \int Z \) has variance zero, so that \( f \int Z = 0 \) almost surely. That is, \( f \int Z = 0 \) almost surely, where \( 1_j \) denotes the indicator function for \( J \). This is true for every Borel set \( J \subseteq K \) so, appealing to the fact that every element of \( H \) is the limit of a sequence of simple functions on \( K \), we have \( \langle Z, b \rangle = 0 \) almost surely for every \( b \in H \). Recall that the covariance operator of \( Z \) is \( \Phi \Sigma \Phi^* \); applying the definition of a covariance operator, we obtain \( \Phi \Sigma \Phi^* (b) = E \langle Z, b \rangle Z \) for all \( b \in H \), so that \( \Phi \Sigma \Phi^* = 0 \). Since \( \Sigma \) is positive definite, this proves that \( \Phi = 0 \). \( \blacksquare \)

With \( \Phi = 0 \), the random-walk component of the Beveridge-Nelson decomposition vanishes, and for each \( t \in \mathbb{N} \) we have \( f_t = w_0 - \hat{\mu}_0 + Ef_t + \hat{\mu}_t \). CKP’s time series of pdfs is thus the sum of a deterministic component \( w_0 - \hat{\mu}_0 \) and \( Ef_t \) (deterministic in the sense of being determined at time zero) and a stationary component \( \hat{\mu}_t \).

The flaw elaborated above with the CKP model derives from the fact that the usual \( L^2 \) Hilbert space is poorly suited to handling pdf-valued random variables, an insight that has been noted in prior mathematical literature. The literature has developed suitable Hilbert spaces in which addition and scalar multiplication are defined such that they preserve the essential properties of pdfs (see, e.g., Egozcue, Díaz-Barrero, and Pawlowsky-Glahn 2006; van den Boogaart, Egozcue, and Pawlowsky-Glahn 2014; Petersen and Müller 2016). I thank Peter Phillips and Won-Ki Seo for pointing me to such references, and for very helpful discussions of how the structures developed therein might form the basis for more useful models of cointegrated pdf-valued time series. Further research in this direction is underway.
References


About the Author

Brendan K. Beare is Associate Professor of Economics at the University of California–San Diego. He has published a range of articles on theoretical and applied econometrics in outlets including *Econ Journal Watch, Econometrica, Econometric Theory*, the *Journal of Applied Econometrics*, and the *Journal of Time Series Analysis*. His email address is bbeare@ucsd.edu.