

# Spurious Regression in Nonstationary Panels with Cross-Unit Cointegration\*

Jean-Pierre Urbain<sup>†</sup> and Joakim Westerlund<sup>‡</sup>

December 13, 2006

## Abstract

This paper illustrates analytically the effects of cross-unit cointegration using as an example the conventional pooled least squares estimate in the spurious panel regression case. The results suggest that the usual result of asymptotic normality depends critically on the absence of cross-unit cointegration.

**JEL Classification:** C13; C33.

**Keywords:** Panel Cointegration; Spurious Regression; Cross-Unit Cointegration; Cross-Section Dependence.

## 1 Introduction

The use of nonstationary panel data techniques to infer both spurious and cointegrated regressions has basically exploded during the last couple of years. The main reason for this being the more powerful and straightforward asymptotic results based on pooling across a large cross-section of independent units. In fact, as shown by Phillips and Moon (1999), since this independence effectively works by smoothing out the usual unit root dependency for each unit, normal inference is usually possible, even in the usually so difficult spurious regression case.

---

\*A preliminary version of the paper was presented at the Conference on Common Feature in Maastricht, at Ente "Luigi Einaudi" in Rome, ESEM2004 in Madrid and at the Econometric Seminar in Rotterdam. The author wants to thank Christian Gengenbach, Stephane Gregoir, Richard Paap, Franco Peracchi and Fabio Buseti for helpful comments and suggestions. The second author would also like to thank the Maastricht Research School of Economics of Technology and Organizations for its hospitality during a visit at the Department of Quantitative Economics at the University of Maastricht, where a part of this paper was written. Thank you also to the Jan Wallander and Tom Hedelius Foundation for financial support under research grant number W2006-0068:1. The usual disclaimer applies.

<sup>†</sup>Corresponding author: Department of Quantitative Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Telephone: +31 43 388 3660; Fax: +31 43 388 48 74; E-mail address: j.urbain@ke.unimaas.nl

<sup>‡</sup>Department of Economics, Lund University, P.O.Box 7082, S-220 07 Lund, Sweden. Telephone: +46 46 222 8670; Fax: +46 46 222 4118; E-mail address: joakim.westerlund@nek.lu.se.

However, as argued by Banerjee *et al.* (2006), assuming independence in this way is usually not warranted, especially when considering macroeconomic or financial data with strong intra-economic linkages. In particular, by using simulation methods, they show that the presence of cross-unit cointegration, a reasonable assumption in most applications, can have quite drastic effects on inference.

In this paper, we complement analytically the simulation results of Banerjee *et al.* (2006) taking as an example the pooled least squares slope estimator in the simple bivariate case.<sup>1</sup> We consider two possibilities, one is when the cointegration occurs across the whole panel, while the other is when the cointegration occurs within groups that are otherwise independent of each other, which seems like a reasonable scenario when studying for example convergence clubs. In both cases, it is shown that the presence of cross-unit cointegration in both the dependent and independent variables leads to statistics that diverge with the size of the cross-section. On the other hand, if at least one of the variables does not cointegrate across units, then normality is again possible, with the center of the distribution located at the long-run average regression coefficient. For the case with cointegration within groups, this naturally leads to the definition of a long-run average group coefficient.

The next two sections present our main findings, while Section 4 concludes. All proofs are relegated to the appendix.

## 2 Panel common unit roots

Suppose that  $y_{it}$  and  $x_{it}$  are two a scalar variables such that

$$x_{it} = \lambda_{xi}f_{xt} + \delta_{xi}e_{xit} + u_{xit} \quad \text{and} \quad y_{it} = \lambda_{yi}f_{yt} + \delta_{yi}e_{yit} + u_{yit}, \quad (1)$$

where  $t = 1, \dots, T$  and  $i = 1, \dots, N$  indexes the time series and cross-sectional units, respectively. Note that while the first term on the right-hand side of these equations are common across  $i$ , the second is not. The parameters  $\lambda_{xi}$ ,  $\delta_{xi}$ ,  $\lambda_{yi}$  and  $\delta_{yi}$  are assumed to be independent of each other, and randomly distributed across  $i$  with expected values  $\lambda_x$ ,  $\delta_x$ ,  $\lambda_y$  and  $\delta_y$ , respectively. We further assume that the disturbances  $u_{xit}$  and  $u_{yit}$  are stationary, while the remaining right-hand side variables in (1) are pure unit root processes that satisfy the following invariance principles as  $T \rightarrow \infty$  for each  $i$

$$\frac{1}{\sqrt{T}} \begin{bmatrix} f_{xt} \\ f_{yt} \end{bmatrix} \Rightarrow \begin{bmatrix} B_x \\ B_y \end{bmatrix} \quad \text{and} \quad \frac{1}{\sqrt{T}} \begin{bmatrix} e_{xit} \\ e_{yit} \end{bmatrix} \Rightarrow \begin{bmatrix} B_{xi} \\ B_{yi} \end{bmatrix},$$

where  $\Rightarrow$  signifies weak convergence. The vectors of common and idiosyncratic Brownian motions,  $B$  and  $B_i$  say, are assumed to be independent with covariance matrices  $\Omega^f$  and  $\Omega_i^e$ , respectively, where  $\Omega_i^e$  is a random matrix with expectation  $\Omega^e$ .

---

<sup>1</sup>This work is related to that of Gengenbach *et al.* (2006). However, while they focus specifically on the case when testing for cointegration, our focus is more conceptual, and our interest in the least squares slope estimate is just for illustration.

Although very simple, as we shall see, this data generating process yields significant insight. In particular, note how (1) nests a variety of spurious regression cases both with and without cross-unit cointegration, depending on the values taken by  $\lambda_{xi}$ ,  $\delta_{xi}$ ,  $\lambda_{yi}$  and  $\delta_{yi}$ .<sup>2</sup> In fact, it is not difficult to see that (1) is actually a system of two Bai and Ng (2004) factor models, one for  $x_{it}$  and one for  $y_{it}$ . Define

$$\widehat{\beta} = \left( \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it} \right) / \left( \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right).$$

The objective of this paper is to show how  $\widehat{\beta}$ , the conventional pooled least squares slope estimator in a regression of  $y_{it}$  on  $x_{it}$ , behaves in this context.

Let us define the following two quantities

$$\begin{aligned} Q_{xx} &= \lambda_x^2 \left( \int B_x^2 \right) + \frac{1}{2} \delta_x^2 \Omega_{xx}^e, \\ U_{xy} &= \frac{1}{6} \left( \lambda_x^2 \delta_y^2 \Omega_{xx}^f \Omega_{yy}^e + \lambda_y^2 \delta_x^2 \Omega_{yy}^f \Omega_{xx}^e + \delta_y^2 \delta_x^2 \Omega_{yy}^e \Omega_{xx}^e \right) \\ &\quad + \frac{1}{2} \lambda_y \lambda_x \delta_y \delta_x \Omega_{xy}^f \Omega_{xy}^e. \end{aligned}$$

Given  $Q_{xx}$  and  $U_{xy}$ , if we let  $\rightarrow_p$  signify convergence in probability, then the asymptotic distribution of  $\widehat{\beta}$  can be summarized in the following way.

**Theorem 1.** Under the assumptions laid out above, as  $N, T \rightarrow \infty$

$$\begin{aligned} \text{(a)} \quad \widehat{\beta} &\rightarrow_p \frac{1}{Q_{xx}} \left( \lambda_y \lambda_x \int B_y B_x + \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e \right) \\ \text{(b)} \quad \sqrt{N} \widehat{\beta} - \sqrt{N} \frac{1}{Q_{xx}} \left( \lambda_y \lambda_x \int B_y B_x + \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e \right) &\Rightarrow N \left( 0, \frac{1}{Q_{xx}^2} U_{xy} \right) \end{aligned}$$

Note how Theorem 1 expresses the asymptotic distribution of  $\widehat{\beta}$  in terms of the parameters of (1). This is very instructive.

Firstly, note that the mean of the distribution depends critically on the common components of  $x_{it}$  and  $y_{it}$ . In particular, note that if  $\lambda_x$  is zero, so that only  $y_{it}$  has a common component, then  $\sqrt{N} \widehat{\beta}$  converges to a normal variate with mean  $\delta_y \Omega_{xy}^e / \delta_x \Omega_{xx}^e$ , which is the long-run average regression coefficient derived by Phillips and Moon (1999) in the cross-sectionally independent case. On the other hand, if  $\lambda_x$  and  $\lambda_y$  are both nonzero, then  $\widehat{\beta}$  is no longer consistent for  $\delta_y \Omega_{xy}^e / \delta_x \Omega_{xx}^e$ , and in fact diverges as  $N$  grows. This suggests that much of the empirical work based on the pooled least squares estimator in nonstationary panel data with a possible common factor structure needs to be reevaluated, see for example Coakley *et al.* (2005).

This clearly illustrates the importance of the common components, both for inferential and cointegration testing purposes. For example, with both  $y_{it}$  and  $x_{it}$  affected, the asymptotic distribution of  $\sqrt{N} \widehat{\beta}$  is no longer normal, but dependent

<sup>2</sup>The case with cointegration between  $y_{it}$  and  $x_{it}$  at is also covered, and amounts to setting  $\lambda_{xi}$  and  $\lambda_{yi}$  to zero for all  $i$  while imposing equality of the idiosyncratic unit roots  $e_{xit}$  and  $e_{yit}$ .

on Brownian motion, which makes it look more like something that one would expect to see in the conventional time series case. Similarly, it is not difficult to see how tests for the absence of cointegration based on the consistency of  $\widehat{\beta}$ , such as those of Kao (1999), become invalid in this case.

Secondly, the variance of the asymptotic distribution of  $\sqrt{N}\widehat{\beta}$  depends on both the idiosyncratic and common components. However, note that if there are no common components, then the variance reduces to  $2\delta_y^2\Omega_{yy}^e/3$ , which is likely to be smaller than for the case with common components, and it is definitely smaller than  $2(\lambda_y^2\Omega_{yy}^f + \delta_y^2\Omega_{yy}^e)/3$ , the variance when  $\lambda_x$  is zero. Hence, the presence of the common components affects the asymptotic distribution of  $\sqrt{N}\widehat{\beta}$  by increasing its variance.

### 3 Group specific unit roots

Next, we examine the case with group specific common unit roots. Specifically, suppose that there are  $M$  groups with  $K$  units in each, and let  $f_{xjt}$  be the common unit root component of  $x_{it}$  for group  $j = 1, \dots, K$  containing the units  $i = 1 + (j - 1)K, \dots, jK$ . The group specific common component of  $y_{it}$  is constructed in the same way, and is denoted  $f_{yjt}$ . Let  $B_{xj}$  and  $B_{yj}$  be the Brownian motions associated with  $f_{xjt}$  and  $f_{yjt}$ , and assume that these have covariance  $\Omega_j^f$  and that they are independent across  $j$ .

In view of Theorem 1, it is not difficult to see that as  $N, T \rightarrow \infty$  with  $M$  fixed

$$\widehat{\beta} \rightarrow_p \left( \lambda_y \lambda_x \frac{1}{M} \sum_{j=1}^M \int B_{yj} B_{xj} + \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e \right) / \left( \lambda_x^2 \frac{1}{M} \sum_{j=1}^M \int B_{xj}^2 + \frac{1}{2} \delta_x^2 \Omega_{xx}^e \right).$$

By contrast, if  $M$  is permitted to grow with  $N$ , then we get

$$\widehat{\beta} \rightarrow_p \left( \lambda_y \lambda_x \Omega_{xy}^f + \delta_x \delta_y \Omega_{xy}^e \right) / \left( \lambda_x^2 \Omega_{xx}^f + \delta_x^2 \Omega_{xx}^e \right),$$

where  $\Omega^f$  is the expected value of  $\Omega_j^f$ . This second result is very interesting. Specifically, suppose that  $\delta_x$  is zero so that  $x_{it}$  has no idiosyncratic unit root, then the probability limit of  $\widehat{\beta}$  reduces to  $\lambda_y \Omega_{xy}^f / \lambda_x \Omega_{xx}^f$ , which might be thought of as a long-run average group coefficient. Thus, if the number of groups are permitted to grow with  $N$ , then we end up with something that is very similar to what we had earlier in the case with panel common unit roots. This seems very reasonable since in this case the group essentially takes the role of the individual unit. Moreover, as in Section 2, it is possible to show that if  $M$  is permitted to grow with  $N$  and  $T$ , then  $\sqrt{N}\widehat{\beta}$  is asymptotically normal.

### 4 Concluding remarks

It is now well-known that the conventional assumption of independent cross-sections is very hard to maintain when conducting inference in nonstationary panels.

In this paper, we study the effects of cross-unit cointegration taking as an example the pooled least squares estimator in the spurious regression case. We consider both the case when the cointegration takes place across the panel as a whole, and when it takes place within groups. Our findings suggest that the presence of cross-unit cointegration can have dramatic effects for the usual asymptotic results based on cross-section independence, and that the consistency and asymptotic normality of the pooled least squares estimator may even be lost. These results have important implications, not only for theoretical, but also for applied work. Take for example the study of Coakley *et al.* (2005), in which the authors seek to infer purchasing power parity using pooled least squares methods in the presence of common factors. Our results indicate that this type of analysis can be potentially very deceptive.

As a final note, although the simplicity of the data generating process used in this paper naturally raises some concerns, we would like to stress that most assumptions are only for ease of exposure, and can in principle be relaxed without affecting the main conclusions. For example, it can be easily generalized to accommodate deterministic intercept, trend terms and additional dynamics.

## References

- Bai, J., and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* 72, 1127-1177.
- Banerjee, A., M. Marcellino and C. Osbat (2004). Some Cautions on the use of Panel Methods for Integrated Series of Macroeconomic Data. *Econometrics Journal* 7, 322-340.
- Coakley, J., R. P. Flood, A. M. Fuertes and M. P. Taylor (2005). Purchasing Power Parity and the Theory of General Relativity: The First Tests. *Journal of International Money and Finance* 24, 293-316.
- Gengenbach, C., F. C. Palm and J-P. Urbain (2006). Panel Cointegration Testing in the Presence of Common Factors. *Oxford Bulletin of Economics and Statistics* 68, 683-719.
- Kao, C. (1999). Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. *Journal of Econometrics* 90, 1-44.
- Phillips, P. C. B., and H. R. Moon (1999). Linear Regression Limit Theory of Nonstationary Panel Data. *Econometrica* 67, 1057-1111.

## Mathematical appendix

In this appendix, Theorem 1 is proven. For brevity, only necessary details are provided. Unreported results can be obtained from the corresponding author upon request.

### Proof of Theorem 1.

Consider (a). The least squares estimator of  $\beta$  can be written as

$$\sqrt{N}\hat{\beta} = \left( \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right)^{-1} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it}. \quad (\text{A1})$$

Consider the denominator, whose limit as  $N, T \rightarrow \infty$  is given by

$$\begin{aligned} \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 &= \lambda_{xi}^2 \frac{1}{T^2} \sum_{t=1}^T f_{xt}^2 + \delta_{xi}^2 \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T e_{xit}^2 \\ &+ 2\lambda_{xi}\delta_{xi} \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T f_{xt} e_{xit} + O_p(T^{-1}) \\ &\rightarrow_p \lambda_x^2 \left( \int B_x^2 \right) + \delta_x^2 E \left( \int B_{ix}^2 \right) + 2\lambda_x \delta_x E \left( \int B_x B_{ix} \right) \\ &= \lambda_x^2 \left( \int B_x^2 \right) + \frac{1}{2} \delta_x^2 \Omega_{xx}^e, \end{aligned} \quad (\text{A2})$$

where the last equality follows from the independence of  $B_x$  and  $B_{ix}$ , and the moments of Brownian motion.

Similarly, for the numerator, we have

$$\begin{aligned} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it} &= \sqrt{N} \lambda_{xi} \lambda_{yi} \frac{1}{T^2} \sum_{t=1}^T f_{xt} f_{yt} + \lambda_{xi} \delta_{yi} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T f_{xt} e_{yit} \\ &+ \lambda_{yi} \delta_{xi} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T f_{yt} e_{xit} \\ &+ \delta_{yi} \delta_{xi} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T e_{yit} e_{xit} + O_p(T^{-1}), \end{aligned} \quad (\text{A3})$$

which has the following limit as  $T \rightarrow \infty$

$$\begin{aligned} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it} &\Rightarrow \sqrt{N} \left( \lambda_{xi} \lambda_{yi} \int B_x B_y \right) \\ &+ \frac{1}{\sqrt{N}} \sum_{i=1}^N \left( \lambda_{xi} \delta_{yi} \int B_x B_{yi} + \lambda_{yi} \delta_{xi} \int B_y B_{xi} + \delta_{yi} \delta_{xi} \int B_{yi} B_{xi} \right) \\ &= \sqrt{N} \left( \lambda_{xi} \lambda_{yi} \int B_x B_y \right) + \frac{1}{\sqrt{N}} \sum_{i=1}^N U_i, \text{ say.} \end{aligned} \quad (\text{A4})$$

Thus, passing  $N \rightarrow \infty$ , it is clear that

$$\frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it} - \left( \lambda_x \lambda_y \int B_x B_y \right) = \frac{1}{N} \sum_{i=1}^N U_i \rightarrow_p \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e.$$

This establishes (a).

Next, consider (b). By using (a) and the Lindberg-Lévy theorem, as  $N, T \rightarrow \infty$

$$\begin{aligned} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it} - \sqrt{N} \left( \lambda_x \lambda_y \int B_x B_y + \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e \right) \\ \Rightarrow N \left( 0, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E(U_i^2) \right). \end{aligned}$$

For the variance, we have

$$\begin{aligned} E(U_i^2) &= E \left( \left( \lambda_x \delta_y \int B_x B_{yi} + \lambda_y \delta_x \int B_y B_{xi} + \delta_y \delta_x \int B_{yi} B_{xi} \right)^2 \right) \\ &= \lambda_x^2 \delta_y^2 E \left( \int B_x B_{yi} \right)^2 + \lambda_y^2 \delta_x^2 E \left( \int B_y B_{xi} \right)^2 + \delta_y^2 \delta_x^2 E \left( \int B_{yi} B_{xi} \right)^2 \\ &\quad + 2 \lambda_y \lambda_x \delta_y \delta_x E \left( \int B_x B_{yi} \right) \left( \int B_{xi} B_y \right) \\ &= \frac{1}{6} \left( \lambda_x^2 \delta_y^2 \Omega_{xx}^f \Omega_{yy}^e + \lambda_y^2 \delta_x^2 \Omega_{yy}^f \Omega_{xx}^e + \delta_y^2 \delta_x^2 \Omega_{yy}^e \Omega_{xx}^e \right) \\ &\quad + \frac{1}{2} \lambda_y \lambda_x \delta_y \delta_x \Omega_{xy}^f \Omega_{xy}^e. \end{aligned} \tag{A5}$$

where the second equality follows from the fact that all cross-products except the squared ones have expectation zero, while the second follows from the moments of Brownian motion.

Putting everything together, we obtain

$$\sqrt{N} \hat{\beta} - \sqrt{N} \frac{1}{Q_{xx}} \left( \lambda_x \lambda_y \int B_x B_y + \frac{1}{2} \delta_x \delta_y \Omega_{xy}^e \right) \Rightarrow N \left( 0, \frac{1}{Q_{xx}^2} U_{xy} \right).$$

This completes the proof of (b). ■