Panel Cointegration Testing in the Presence of Linear Time Trends

Uwe Hassler

with Mehdi Hosseinkouchack
Goethe-Universität Frankfurt

20^{th} IPDC, Tokyo, July 9, 2014

Moving on to ...

Empirical motivation

- Many nonstationary panels are likely to be driven by linear time trends, see the applications by ...
- Coe and Helpman (1995) (and Westerlund (2005a)) on R&D spillovers (total factor productivity and capital stock),
- Larsson, Lyhagen and Löthgren (2001) on log. real consumption and dis. income (per capita), and inflation,
- Westerlund and Edgerton (2008) on exchange rates and price differential,
- Hanck (2009) on prices (weak PPP).

Health care expenditure and GDP

- A sequence of papers published in the *Journal of Health Economics* addresses the issue of linear time trends:
- McCoskey and Selden (1998) - ignoring linear time trends,
- Hansen and King (1998),
- Blomqvist and Carter (1997), detrended regressions,
- Gerdtham and Löthgren (2000), (partly) detrended, regressions,
- Westerlund (2007, OxBull), detrended regressions

How to handle a linear time trend?

- Many papers propose detrending of the data; see e.g. Breitung (2005), Chang and Nguyen (2011), Karaman Örsal and Droge (2013), Demetrescu, Hanck, and Tarcolea (2014); more below;
- empirically relevant often: explain one trend by another *without* detrending (to increase power and for economic reasons).
- The latter case addressed only by Kao (1999), but not completely
- This is where the present paper comes in.

Our framework: What we do

- We consider the following class: Single-equation approaches building on OLS...
- that test for the null of either cointegration or no cointegration...
- and are residual-based or not.

Beyond our framework

- Most panel cointegration tests in single-equations setting, notable exceptions being Larsson, Lyhagen and Löthgren (2001), Groen and Kleibergen (2003), Breitung (2005) and Karaman Örsal and Droge (2013).
- Recent single-equation tests by Chang and Nguyen (2012) or Demetrescu, Hanck and Tarcolea (2014) rely on nonlinear instrumental variable estimation.

Summary

- If at least one regressor is dominated by a linear time trend ...
- ... then limiting distributions and critical values provided for and applied with the situation "with intercept only" are not correct
- ... and their usage results in size distortions growing with the panel size N, while correct critical value are available from the literature:
- Regression on k I(1) variables with drift and on *intercept only* amounts to limiting distribution arising from regression on $k-1$ I(1) variables and *intercept plus a linear time trend*.

Outline

Moving on to ...

Assumption 1: I(1) with drift

- Partition the m-vector $z_{i,t}$ of observables into a scalar $y_{i,t}$ and a k-element vector $x_{i,t}$, $z'_{i,t} = (y_{i,t}, x'_{i,t})$, $m = k + 1$.
- Allow for linear time trends (drift), $i = 1, \ldots, N$:

$$z_{i,t} = \mu_i \, t + \sum_{j=1}^{t} e_{i,j} = \begin{pmatrix} \mu_{i,y} \\ \mu_{i,x} \end{pmatrix} t + \sum_{j=1}^{t} \begin{pmatrix} e_{i,y,j} \\ e_{i,x,j} \end{pmatrix}, \quad t = 1, \ldots, T .$$

- where the vector $\{x_{i,t}\}$ alone is not cointegrated,
- $\{z_{i,t}\}$ may be cointegrated or not.

Assumption 2: Individual tests

- Let $\bar{S}_i^{(m)}$ and $\widetilde{S}_i^{(m)}$ stand for statistics computed from regressions with "intercept only" and "intercept plus trend", respectively.
- We assume limiting distributions free of nuisance parameters under the null,

$$\begin{aligned} \bar{S}_i^{(m)} &\Rightarrow \bar{\mathcal{L}}_i^{(m)} \quad \text{if } \mu_{i,x} = 0\,, \\ \widetilde{S}_i^{(m)} &\Rightarrow \widetilde{\mathcal{L}}_i^{(m)} \quad \text{for all } \mu_{i,x}\,, \end{aligned}$$

- as $T \to \infty$

Panel statistics for $H_0 = \bigcap_{i=1}^{N} H_{i,0}$

- Group statistics:

$$\bar{G}^{(m)} = \frac{1}{N} \sum_{i=1}^{N} \bar{S}_i^{(m)} \quad \text{or} \quad \widetilde{G}^{(m)} = \frac{1}{N} \sum_{i=1}^{N} \widetilde{S}_i^{(m)} .$$

- Pooled statistics: $\bar{P}^{(m)}$ or $\widetilde{P}^{(m)}$ equal

$$g\left(\sum_{i=1}^{N} \bar{N}_{i,T}^{(m)} , \sum_{i=1}^{N} \bar{D}_{i,T}^{(m)} \right) \text{ or } g\left(\sum_{i=1}^{N} \widetilde{N}_{i,T}^{(m)} , \sum_{i=1}^{N} \widetilde{D}_{i,T}^{(m)} \right) .$$

- Combination of p values: not today

Assumption 3: Panel tests

- Let $\bar{Z}^{(m)}$ and $\widetilde{Z}^{(m)}$ stand for $\bar{G}^{(m)}$ and $\widetilde{G}^{(m)}$ or for $\bar{P}^{(m)}$ and $\widetilde{P}^{(m)}$, respectively
- Let under the null hold as $T \to \infty$ followed by $N \to \infty$

$$\begin{aligned} \sqrt{N}\left(\bar{Z}^{(m)} - \bar{\mu}_m\right) &\Rightarrow \mathcal{N}(0, \bar{\sigma}_m^2) \quad \text{if } \mu_{i,x} = 0\,, \\ \sqrt{N}\left(\widetilde{Z}^{(m)} - \widetilde{\mu}_m\right) &\Rightarrow \mathcal{N}(0, \widetilde{\sigma}_m^2) \quad \text{for all } \mu_{i,x}\,. \end{aligned}$$

Moving on to ...

Kao (1999)

- Kao (1999) is the only panel paper addressing regressions with intercept only under linear time trends
- For a residual-based test for no cointegration Kao (1999, Theo. 4) claimed for a pooled statistic

$$\sqrt{N}\left(\bar{P}^{(m)} - \widetilde{\mu}_1\right) \Rightarrow \mathcal{N}(0, \widetilde{\sigma}_1^2) \tag{1}$$

- We find, however, that this is correct only for $m = 2$

Theorem 1

- Let $\mu_{i,x} \neq 0$, $i = 1, \ldots, N$. Then it holds under the above assumptions:

a) $\bar{S}_i^{(m)} \Rightarrow \widetilde{\mathcal{L}}^{(m-1)}$ as $T \to \infty$;

b) $\sqrt{N}\left(\bar{Z}^{(m)} - \widetilde{\mu}_{m-1}\right) \Rightarrow \mathcal{N}(0, \widetilde{\sigma}^2_{m-1})$ as $N \to \infty$.

2 strategies

- S_I: Ignore possibility of linear trend when working with intercept only: Compare $\bar{S}_i^{(m)}$ with $\bar{\mathcal{L}}^{(m)}$, and $\bar{G}^{(m)}$ or $\bar{P}^{(m)}$ with $\bar{\mu}_m$ and $\bar{\sigma}_m$.
- S_A: Always account for possibility of linear trend when working with intercept only: Compare $\bar{S}_i^{(m)}$ with $\widetilde{\mathcal{L}}^{(m-1)}$, and $\bar{G}^{(m)}$ or $\bar{P}^{(m)}$ with $\widetilde{\mu}_{m-1}$ and $\widetilde{\sigma}_{m-1}$.

Examples

- Pedroni (1999, 2004): Residual-based test for no cointegration (Dickey-Fuller statistics)
- Westerlund (2005): Residual-based test for no cointegration (Breitung statistic)
- Westerlund (2007): Error-correction test for no cointegration (t statistic)
- Westerlund (2005a): Residual-based test for cointegration (CUSUM statistic)

All those tests satisfy the following corollary.

Corollary 1: Further assume $\widetilde{\mu}_{m-1} < \bar{\mu}_m$

Under the null hypothesis one has the following.

a) For a test rejecting for too negative values, the probability to reject ...

$$\ldots \begin{array}{c} \text{increases with growing } N \text{ to 1 under } S_I \text{ if } \mu_{i,x} \neq 0 \\ \text{decreases with growing } N \text{ to 0 under } S_A \text{ if } \mu_{i,x} = 0 \end{array} ;$$

b) for a test rejecting for too large values, the probability to reject ...

$$\ldots \begin{array}{c} \text{decreases with growing } N \text{ to 0 under } S_I \text{ if } \mu_{i,x} \neq 0 \\ \text{increases with growing } N \text{ to 1 under } S_A \text{ if } \mu_{i,x} = 0 \end{array} .$$

Approximate effective size of the group t-test by Pedroni under S_l for $\mu_{i,x} \neq 0$

	$N =$	10	20	30	40	50
	$\alpha = 0.01$	0.030	0.053	0.079	0.107	0.137
$k = 1$	$\alpha = 0.05$	0.126	0.190	0.249	0.307	0.361
	$\alpha = 0.10$	0.227	0.314	0.389	0.455	0.515
	$\alpha = 0.01$	0.017	0.024	0.030	0.036	0.043
$k = 2$	$\alpha = 0.05$	0.080	0.102	0.122	0.141	0.159
	$\alpha = 0.10$	0.154	0.188	0.217	0.243	0.268
	$\alpha = 0.01$	0.014	0.017	0.020	0.022	0.025
$k = 3$	$\alpha = 0.05$	0.067	0.078	0.087	0.096	0.104
	$\alpha = 0.10$	0.130	0.148	0.162	0.175	0.187

Approximate effective size of the group t-test by Pedroni under S_A for $\mu_{i,x} = 0$

	$N =$	10	20	30	40	50
	$\alpha = 0.01$	0.003	0.001	0.001	0.000	0.000
$k = 1$	$\alpha = 0.05$	0.018	0.009	0.006	0.004	0.002
	$\alpha = 0.10$	0.038	0.022	0.014	0.009	0.006
	$\alpha = 0.01$	0.006	0.004	0.003	0.002	0.002
$k = 2$	$\alpha = 0.05$	0.030	0.023	0.018	0.014	0.012
	$\alpha = 0.10$	0.063	0.049	0.040	0.033	0.028
	$\alpha = 0.01$	0.007	0.006	0.005	0.004	0.004
$k = 3$	$\alpha = 0.05$	0.037	0.031	0.027	0.024	0.022
	$\alpha = 0.10$	0.076	0.065	0.058	0.053	0.048

Phillips-Ouliaris test

- Consider

$$y_t = \bar{c} + \bar{\beta}' x_t + \bar{u}_t \quad \text{or} \quad y_t = \widetilde{c} + \widetilde{\delta}\, t + \widetilde{\beta}' x_t + \widetilde{u}_t \qquad (2)$$

  under the null hypothesis of no cointegration
- Phillips and Ouliaris (1990) establish for residual-based Dickey-Fuller tests Assumption 2.
- Hansen (1992a) proves Theorem 1 a, but observes that critical values from $\bar{\mathcal{L}}^{(m)}$ and $\widetilde{\mathcal{L}}^{(m-1)}$ are (coincidentally) almost identical.

The 2 strategies for Phillips-Ouliaris-Hansen

- S_I: Apply $\bar{\mathcal{L}}^{(m)}$ in the case of "intercept only": mildly liberal under the null if $\mu_x \neq 0$.
- S_A: Account for the possibility of linear trends by always applying $\widetilde{\mathcal{L}}^{(m-1)}$ in the case of "intercept only": mildly conservative under the null if $\mu_x = 0$.
- Hansen (1992a) advocated S_A

Further examples

Theorem 1 a) also applies to ...

- Parameter stability test (no cointegration) by Hansen (1992b)
- Residual-based fluctuation cointegration test by Xiao (1999)
- Residual-based CUSUM cointegration test by Xiao and Phillips (2002)

... where strategies S_I and S_A have different consequences (qualitatively and quantitatively)

Combination of p-values

- Compute individual p-values p_i, $i = 1, \ldots, N$,
- Maddala and Wu (1999) and Choi (2001) suggest classical combination methods (Fisher or inverse normal) for *independent* individuals
- Corrections for *cross-dependence* discussed in Hartung (1999), see also Demetrescu, Hassler, Tarcolea (2006) and Hanck (2009),
- Distortions observed in the pure time series case grow fast with N

Moving on to ...

Experimental size of the group t-test by Westerlund (2007) under S_l for $\mu_{i,x} \neq 0$

	$N =$	10	20	30	40	50	100
	$\alpha = 0.01$	0.129	0.335	0.527	0.693	0.811	0.991
$k = 1$	$\alpha = 0.05$	0.367	0.635	0.804	0.905	0.951	0.999
	$\alpha = 0.10$	0.539	0.781	0.900	0.960	0.982	1.000
	$\alpha = 0.01$	0.082	0.180	0.301	0.415	0.534	0.886
$k = 2$	$\alpha = 0.05$	0.267	0.444	0.596	0.707	0.798	0.976
	$\alpha = 0.10$	0.412	0.605	0.747	0.830	0.896	0.992
	$\alpha = 0.01$	0.069	0.151	0.246	0.325	0.431	0.786
$k = 3$	$\alpha = 0.05$	0.227	0.382	0.519	0.614	0.709	0.939
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Experimental size of the group t-test by Westerlund (2007) under S_A for $\mu_{i,x} \neq 0$

	$N =$	10	20	30	40	50	100
	$\alpha = 0.01$	0.010	0.009	0.008	0.008	0.008	0.006
$k = 1$	$\alpha = 0.05$	0.048	0.048	0.043	0.040	0.041	0.036
	$\alpha = 0.10$	0.093	0.095	0.087	0.080	0.083	0.069
	$\alpha = 0.01$	0.008	0.009	0.007	0.009	0.006	0.005
$k = 2$	$\alpha = 0.05$	0.046	0.040	0.039	0.040	0.034	0.031
	$\alpha = 0.10$	0.093	0.082	0.083	0.082	0.073	0.066
	$\alpha = 0.01$	0.010	0.010	0.012	0.010	0.011	0.011
$k = 3$	$\alpha = 0.05$	0.051	0.052	0.052	0.049	0.052	0.053

Experimental size of the group t-test by Westerlund (2007) under detrending

	$N =$	10	20	25	50	100
	$\alpha = 0.01$	0.009	0.008	0.009	0.008	0.008
$k = 1$	$\alpha = 0.05$	0.047	0.045	0.044	0.042	0.039
	$\alpha = 0.10$	0.094	0.091	0.089	0.085	0.080
	$\alpha = 0.01$	0.010	0.013	0.012	0.014	0.015
$k = 2$	$\alpha = 0.05$	0.053	0.059	0.057	0.058	0.059
	$\alpha = 0.10$	0.106	0.113	0.108	0.107	0.111
	$\alpha = 0.01$	0.011	0.012	0.011	0.011	0.009
$k = 3$	$\alpha = 0.05$	0.050	0.054	0.050	0.051	0.047
	$\alpha = 0.10$	0.097	0.103	0.104	0.098	0.095

Experimental power of the group t-test by Westerlund (2007) for $\mu_{i,x} \neq 0$ at 5% size

Under S_A

$N =$	10	20	25	50	100
$k = 1$	0.495	0.726	0.802	0.965	1.000
$k = 2$	0.281	0.437	0.508	0.742	0.942
$k = 3$	0.157	0.224	0.251	0.394	0.620
detrended regression					
$k = 1$	0.235	0.398	0.453	0.691	0.922
$k = 2$	0.169	0.252	0.291	0.454	0.714
$k = 3$	0.092	0.114	0.124	0.183	0.271

Moving on to ...

Work in progress – things to do

- More Monte Carlo
- Empirical application.
- Further "strategies" in addition to S_I or S_A: (i) always detrending: expected power losses; (ii) pretesting on $\mu_x = 0$: will affect subsequent inference; (iii) combine evidence from S_I AND S_A: "cointegration testing under uncertainty about linear trends"