Mr Francesco Devere Bispham
University of Hull
Department of Economics
Hull HU6 7RX
U.K.
F.D.Bispham@econ.hull.ac.uk*

Bootstrap Confidence Intervals for a Panel Data Cointegration Regression

Abstract

In this paper we construct bootstrap confidence intervals for a panel data cointegration regression. The bootstrap is shown to be at least as efficient as first-order asymptotic theory, when it is used with the Dynamic OLS (DOLS) estimators proposed by Kao and Chiang (2000) and Pedroni (2001) and a large macro panel of 20 OECD countries, for quarterly observations from 1957Q1-1991Q2. Using the optimal confidence interval criterion the errors in coverage probabilities are shown to be smallest with the bootstrap and it also produces the shortest intervals.

Keywords: Bootstrap, dynamic OLS estimator, efficient estimation, panel cointegration regression

JEL Classification: C13, C14, C15, C16, C22, C23, C32, C33

*Acknowledgments: I would gratefully like to thank Jurgen Doornik, Peter Phillips and Karim Abadir for their helpful comments and suggestions. I would also like to thank Keshab Bhattarai, Allan Reese and all the staff at Hull University computing centre

1 Introduction

Since its introduction by Efron (1979) the bootstrap has been the focus of much research in statistics and econometrics. Numerous books have appeared on the topic, eg. Beran and Ducharne (1991), Davidson and Hinkley (1997), Efron and Tibshirani (1993) and Hall (1992). Some bootstrap research papers with an econometric orientation are Maddalla and Jeong (1996), Hall (1994), Horowitz (1997) and Vinod (1993). The bootstrap is a method by which one can estimate the distribution of an estimator or test statistic by resampling ones data. One actually treats the data as if it were the population for the purpose of evaluating the distribution of interest. In finite samples the bootstrap is often more accurate than first-order asymptotic approximations. Thus it is a practical method of improving upon first-order asymptotic approximations. Such improvements are called asymptotic refinements and lead in general to more efficient estimation. The bootstrap can provide asymptotic refinements in a number of situations, eg hypothesis testing and confidence interval estimation. The bootstrap can be used to obtain confidence intervals with reduced errors in coverage probabilities. That is the difference between the true and nominal coverage probabilities is often lower when the bootstrap is used than when first-order asymptotic approximations are used to obtain a confidence interval.
This paper aims to show how a number of bootstrap procedures can be used to construct confidence intervals for a panel cointegration regression estimated by Dynamic Ordinary Least Squares (DOLS) procedures. Moreover it shows how use of the bootstrap leads to more efficient estimation of regression parameters and confidence intervals. The layout of the paper is as follows. In section 2 we show the bootstrap methodology. In section 3 the bootstrap procedures are given. In sections 4 and 5 the Dynamic OLS estimation of a panel cointegration regression is described. In section 6 efficient estimation is discussed. In section 7 we have the estimation results and conclusion.

2 The Bootstrap Methodology

We now give a general method, without proofs, of how the bootstrap is implimented and how it can be used to construct confidence intervals.

Definition 2.1 *Let the data $\{X_i, i = 1, \ldots, N\}$ be a random sample from a probability distribution with cumulative distribution function (CDF) F_0. Let $F_0 \in \vartheta$ where ϑ is a finite or infinite family of distribution functions. We* *may index a parameter by θ whose population value is θ_0 and write $F_0(X, \theta_0)$ for $P(X \leq x)$.*

Definition 2.2 *Let $T_n = T_n(X_i, \ldots X_n)$ be a statistic (a function of the data) and let $G_n(\tau, F_0) = P(T_n \leq \tau)$ be the exact finite sample CDF of T_n when the data are sampled from the distribution of F_0. If $G_n(., F_0)$ does not depend on F_0 the statistic T_n is then said to be* **pivotal**. *An* **asymptotically pivotal** *statistic means that the asymptotic distribution does not depend on unknown population parameters.*

Remark 2.3 *The t-statistic for testing a hypothesis about a slope coefficient in a normal linar regression model is independent of unknown population parameters and therefore is pivotal.*

The bootstrap is a method for estimating $G_n(., F_0)$ or features of $G_n(., F_0)$ when F_0 is unknown and is usually implimented by Monte Carlo simulation. Asymptotic distribution theory is another method for estimating $G_n(., F_0)$. Many econometric statistics are asymptotically standard normally distributed, possibly after centering and normalisation. Such statistics are asymptotically pivotal.

Example 2.4 Monte Carlo procedure for Bootstrap estimation of $\mathbf{G}_n(\tau, F_0)$

(a) *Step 1: Generate a bootstrap sample of size n, $\{X_i^*, i = 1, \ldots, N\}$ by sampling the distribution corresponding to F_n randomly. If F_n is the empirical distribution function (EDF) of the estimation data set then, the bootstrap sample can be obtained by sampling the estimation data randomly with replacement.*

(b) *Step 2: Compute $T_n^* = T_n(X_i^*, \ldots, X_n^*)$.*

(c) *Step 3: Use the results of many repetitions of steps 1 and 2 to compute the empirical probability of the event $T_n^* \leq \tau$, ie $P^*(T_n^* \leq \tau)$.*

Proposition 2.5 *Let $G_\infty(., F_0)$ denote the asymptotic CDF of T_n when the data are sampled from the distribution whose CDF is F_0. If T_n is asymptotically pivotal then $G_\infty(., F_0) = G_\infty(.)$ does not depend on F_0. Therefore, when n is sufficiently large $G_n(., F_0)$ can be estimated by $G_\infty(.)$ without knowing F_0.*

Theorem 2.6 *Let θ be a population parameter with unknown but true value θ_0. Let θ_n be a $\sqrt{n}$-consistent, asymptotically normal estimator of θ and let s_n be a consistent estimator of the standard deviation of the asymptotic distribution of $\sqrt{n}(\theta_n - \theta_0)$. Then an* **asymptotic** $(1-\alpha)$ **confidence interval** *for θ_0 is*

$$\theta_n - z_{(\infty,\frac{\alpha}{2})}\frac{s_n}{\sqrt{n}} \leq \theta_0 \leq \theta_n + z_{(\infty,\frac{\alpha}{2})}\frac{s_n}{\sqrt{n}}. \tag{1}$$

Corollary 2.7 *Define $T_n = \frac{\sqrt{n}(\theta_n - \theta_0)}{s_n}$. Then the coverage probability of the asymptotic confidence interval is*

$$P(|T_n| \leq z_{(\infty,\frac{\alpha}{2})}). \tag{2}$$

Remark 2.8 *If $z_{(n,\frac{\alpha}{2})}$ is the exact, finite sample α-level critical value, then it is the $(1-\alpha)$ quantile of the distribution of T_n. Therefore $z_{(n,\frac{\alpha}{2})}$ solves the equation*

$$G_n(z_{(n,\frac{\alpha}{2})}, F_0) - G_n(-z_{(n,\frac{\alpha}{2})}, F_0) = 1 - \alpha. \tag{3}$$

Unless T_n is exactly pivotal we cannot solve the above since F_0 is unknown. A feasible version can be obtained from a first-order asymptotic approximation by replacing G_n by G_∞. Thus the asymptotic critical value $z_{(\infty,\frac{\alpha}{2})}$ solves

$$G_\infty(z_{(\infty,\frac{\alpha}{2})}, F_0) - G_\infty(-z_{(\infty,\frac{\alpha}{2})}, F_0) = 1 - \alpha. \tag{4}$$

When T_n is asymptotically pivotal $G_\infty(.)$ is the standard normal distribution and $z_{(\infty,\frac{\alpha}{2})}$ can be obtained from the standard normal quantiles. The bootstrap version of Theorem (2.6) replaces F_0 with the EDF of the data F_n. The bootstrap critical value $z^*_{(n,\frac{\alpha}{2})}$ solves

$$G_n(z^*_{(n,\frac{\alpha}{2})}, F_n) - G_n(-z^*_{(n,\frac{\alpha}{2})}, F_n) = 1 - \alpha. \tag{5}$$

Theorem 2.9 *Let θ^*_n be our bootstrap estimator of θ_n and s^*_n be our bootstrap estimator of s_n thus $T^*_n = \frac{\sqrt{n}(\theta^*_n - \theta_0)}{s^*_n}$ is our bootstrap statistic computed by the repetitions of steps 1-2 of our Monte Carlo procedure for bootstrap estimation. Let these repetitions be the empirical distribution of $|T_n|$ and let $z^*_{(n,\frac{\alpha}{2})}$ equal the $(1-\alpha)$ quantile of the distribution satisfying equation (5). Then the* **bootstrap** $(1-\alpha)$ **confidence interval** *for θ_0 is*

$$\theta_n - z^*_{(n,\frac{\alpha}{2})}\frac{s_n}{\sqrt{n}} \leq \theta_0 \leq \theta_n + z^*_{(n,\frac{\alpha}{2})}\frac{s_n}{\sqrt{n}}. \tag{6}$$

Corollary 2.10 *The coverage probability of the bootstrap confidence interval is*

$$P(|T_n| \leq z^*_{(n,\frac{\alpha}{2})}). \tag{7}$$

3 Bootstrap Procedures

(A) The standard or asymptotic confidence interval[1]

The interval estimate for a parameter β_k is just as useful as a point estimate. Together they tell us what the best estimate for β_k is and how much error we can expect. Texts such as Cramer (1946), Malinvaud (1980) and Casella and Berger (2002) provide general discussions of the interval estimate. Large sample theory is often used here with unknown confidence interval parameters substituted by their large sample plug-in estimates, which then provides asymptotic justification for the confidence interval.

Assume $Z = \frac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \sim N(0,1)$ and let $z^{(\alpha)}$ be the 100αth percentile point of a $N(0,1)$ distribution as given by the standard normal table. Thus for $\alpha = 0.025$ and 0.05 then $z^{(0.025)} = -1.96$ and $z^{(0.05)} = -1.645$ and $z^{(1-\alpha)} = z^{(0.975)} = 1.96$ and $z^{(0.95)} = 1.645$, respectively etc. Thus we can write

$$Prob\left\{z^{(\alpha)} \leq \frac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \leq z^{(1-\alpha)}\right\} = 1 - 2\alpha \tag{8}$$

or

$$Prob\left\{\hat{\beta}_k - z^{(1-\alpha)} se(\hat{\beta}_k) \leq \beta_k \leq \hat{\beta}_k - z^{(\alpha)} se(\hat{\beta}_k)\right\} = 1 - 2\alpha \tag{9}$$

or

$$Prob\left\{\beta_k \in \left[\hat{\beta}_k - z^{(1-\alpha)} se(\hat{\beta}_k), \hat{\beta}_k - z^{(\alpha)} se(\hat{\beta}_k)\right]\right\} = 1 - 2\alpha. \tag{10}$$

In general we can write

$$\left[\hat{\beta}_k - z^{(1-\alpha)} se(\hat{\beta}_k), \hat{\beta}_k - z^{(\alpha)} se(\hat{\beta}_k)\right] \tag{11}$$

as the standard confidence interval for β_k with coverage probability $= 1 - 2\alpha$. We can also write the confidence interval as

$$\left[\hat{\beta}_k \pm z^{(1-\alpha)} se(\hat{\beta}_k)\right]. \tag{12}$$

The latter formula showing that $z^{(\alpha)} = -z^{(1-\alpha)}$ which when $\alpha = 0.05$ and $1 - 2\alpha = 0.90$ we get

$$\left[\hat{\beta}_k \pm 1.645 se(\hat{\beta}_k)\right]. \tag{13}$$

[1] When constructing confidence intervals for multiparameter vectors, eg $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)'$ we get k-dimensional confidence rectangles. To avoid notational difficulties we shall restrict our bootstrap confidence intervals to the single parameter case, ie the element $\hat{\beta}_k$ of $\hat{\beta}$. However, it should be noted that these single parameters belong to multiparameter vectors

We can also write equations (11),(12) and (13) in terms of the upper and lower confidence bounds, that is

$$\hat{\theta}^s_{lo} = \left[\hat{\beta}_k - z^{(1-\alpha)} se(\hat{\beta}_k)\right] \quad = \text{Lower Bound.} \tag{14}$$

$$\hat{\theta}^s_{up} = \left[\hat{\beta}_k - z^{(\alpha)} se(\hat{\beta}_k)\right] \quad = \text{Upper Bound.} \tag{15}$$

Hence the standard $(1 - 2\alpha)$ confidence interval becomes $\left[\hat{\theta}^s_{lo}, \hat{\theta}^s_{up}\right]$.
The above confidence intervals are exact. However assuming $Z = \frac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \sim N(0,1)$ holds only asymptotically, then these confidence intervals become approximations with large sample justification only.

(B) The Percentile Method

This was developed by Efron (1981),(1982) and uses the bootstrap estimates of β_k to construct a confidence interval. Hall (1992) also develops a percentile confidence interval and refers to Efron's as the, "other" percentile method. Given the bootstrap data set $\{X^{*b}_{it}, i = 1, \ldots, N, t = 1, \ldots, T\}$ for $b = 1, \ldots, B$, let the vector of bootstrap replications (ie $\hat{\beta}^*(b) = s(X^{*b}_{it})$, the estimate of β_k), be $\hat{\beta}^*$. Let $\hat{G}$ be the cumulative distribution function (CDF) of $\hat{\beta}^*$. Then the exact $(1 - 2\alpha)$ percentile confidence interval is defined by the α and $(1 - \alpha)$ percentiles of $\hat{G}$

$$\left[\hat{\theta}^p_{lo}, \hat{\theta}^p_{up}\right] = \left[\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)\right]. \tag{16}$$

Since $\hat{G}^{-1}(\alpha) = \hat{\beta}^{*(\alpha)} = 100\alpha$th percentile of the bootstrap distribution and $\hat{G}^{-1}(1-\alpha) = \hat{\beta}^{*(1-\alpha)} = 100(1-\alpha)$th percentile of the bootstrap distribution we have

$$\left[\hat{\theta}^p_{lo}, \hat{\theta}^p_{up}\right] = \left[\hat{\beta}^{*(\alpha)}, \hat{\beta}^{*(1-\alpha)}\right]. \tag{17}$$

To impliment this in practice one uses a finite number of bootstrap replications. It is well known that the number of replications required to compute a confidence interval is around 1000 and is much greater than the number required to compute standard errors, ie around 100 (see Hall (1986) on the number of bootstrap replications needed to form a confidence interval). The percentile method does, however, have problems when used with small samples or with asymmetric distributions.

Percentile Bootstrap Algorithm

1. Generate B independent bootstrap data sets $X^{*1}, X^{*2}, \ldots, X^{*B}$, where $X^{*j} = \{X^{*j}_{it}, i = 1, \ldots, N, t = 1, \ldots, T\}$.

2. Compute the bootstrap replication $\hat{\beta}^*(b) = s(X_{it}^{*b})$ for $b = 1, \ldots, B$.

3. Let $\hat{\beta}_B^{*(\alpha)}$ be the 100αth empirical percentile of the $\hat{\beta}^*(b)$ values, ie the B.αth value in the ordered list of B replications.

Hence if $B = 2000$ and $\alpha = 0.05$ then $\hat{\beta}_B^{*(\alpha)}$ is the 100th ordered value of the replications. Similarly $\hat{\beta}_B^{*(1-\alpha)}$ is the $100(1-\alpha)$th empirical percentile.
Thus the approximate $(1-2\alpha)$ percentile interval is

$$\left[\hat{\theta}_{lo}^{p}, \hat{\theta}_{up}^{p}\right] = \left[\hat{\beta}_B^{*(\alpha)}, \hat{\beta}_B^{*(1-\alpha)}\right]. \tag{18}$$

B here denotes that the approximation is based on B replications. As $B \to \infty$ then $\hat{\beta}_B^{*(\alpha)} \to \hat{\beta}^{*(\alpha)}$ and $\hat{\beta}_B^{*(1-\alpha)} \to \hat{\beta}^{*(1-\alpha)}$.

(C) Bias Corrected Method (BC)

Both the bias corrected and bias corrected and accelerated methods modify the percentile bootstrap method. Both were introduced by Efron (1987) and Efron and Tibsharani (1986). They depend on two parameters (i) $\hat{a}$ called the acceleration and (ii) $\hat{z}_0$ called the bias correction. We now show how to calculate the bias corrected interval endpoints

$$\left[\hat{\theta}_{lo}^{bc}, \hat{\theta}_{up}^{bc}\right] = \left[\hat{\beta}^{*(\alpha_1)}, \hat{\beta}^{*(\alpha_2)}\right] \tag{19}$$

where $\alpha_1 = \Phi\left(2\hat{z}_0 + z^{(\alpha)}\right)$ and $\alpha_2 = \Phi\left(2\hat{z}_0 + z^{(1-\alpha)}\right)$.
Hence

$$\left[\hat{\theta}_{lo}^{bc}, \hat{\theta}_{up}^{bc}\right] = \left[\hat{G}^{-1}(\alpha_1), \hat{G}^{-1}(\alpha_2)\right], \tag{20}$$

$$= \left[\hat{G}^{-1}\left(\Phi\left(2\hat{z}_0 + z^{(\alpha)}\right)\right), \hat{G}^{-1}\left(\Phi\left(2\hat{z}_0 + z^{(1-\alpha)}\right)\right)\right]. \tag{21}$$

Where $\Phi(.)$ is the standard normal cumulative distribution function and $z^{(\alpha)}$ is the 100αth percentile point of a standard normal distribution, eg $z^{(0.95)} = 1.645$ and $\Phi(1.645) = 0.95$. When $\hat{z}_0 = 0$ the BC interval is the same as the percentile interval.

Computation of $\hat{z}_0$

We compute $\hat{z}_0$ directly from the proportion of bootstrap replications less than the original estimate of β_k, ie $\hat{\beta}_k$,

$$\hat{z}_0 = \Phi^{-1}\left(\# \frac{\{\hat{\beta}^*(b) < \hat{\beta}_k\}}{B}\right). \tag{22}$$

Where Φ^{-1} is the inverse function of a standard normal cumulative distribution function, eg $\Phi^{-1}(0.95) = 1.645$.

(D) Bias Corrected and Accelerated Method (BC_a)

Here we calculate the BC_a interval endpoints as

$$\left[\hat{\theta}_{lo}^{bc_a}, \hat{\theta}_{up}^{bc_a}\right] = \left[\hat{\beta}^{*(\alpha_1)}, \hat{\beta}^{*(\alpha_2)}\right] \tag{23}$$

where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})}\right) \tag{24}$$

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})}\right). \tag{25}$$

Hence

$$\left[\hat{\theta}_{lo}^{bc_a}, \hat{\theta}_{up}^{bc_a}\right] = \left[\hat{G}^{-1}(\alpha_1), \hat{G}^{-1}(\alpha_2)\right] \tag{26}$$

$$= \left[\hat{G}^{-1}\left(\Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})}\right)\right), \hat{G}^{-1}\left(\Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})}\right)\right)\right]. \tag{27}$$

Where $\Phi(.)$ is as before. Again when $\hat{a} = \hat{z}_0 = 0$ the BC_a interval is the same as the percentile interval.

Computation of $\hat{a}$

Of the number of parametric and non-parametric ways to compute $\hat{a}$, the simplest is probably the Jackknife estimate which we now explain. The jackknife method was developed by Quenouille (1949) and discussed in Efron (1982). See Wu (1986) for applications to regression analysis.

The Delete-One Jackknife in the Panel Data Regression Model

Given the linear panel data model

$$y_{it} = \alpha_i + x_{it}'\beta + e_{it} \tag{28}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$,
where $\{y_{it}\} \sim I(1), \{x_{it}\} \sim I(1)$, are random variables, $\{e_{it}\} \sim I(0)$ a stationary disturbance term, and β and α_i are $((k-1)\times 1)$ and (1×1) parameters of interest, respectively. Then writing in matrix form

$$y = [I_N \otimes i_T]\beta_1 + X_s\beta_s + e \tag{29}$$

$$= [I_N \otimes i_T \quad X_s] \quad \begin{pmatrix} \beta_1 \\ \beta_s \end{pmatrix} + e \tag{30}$$

$$= X\beta + e \tag{31}$$

where $y = (y_{11}, \ldots, y_{NT})'$ is $(NT \times 1)$, $e = (e_{11}, \ldots, e_{NT})'$ is $(NT \times 1)$, $i_T = (1, 1, \ldots, 1)'$ is $(T \times 1)$, $\beta_s = (\beta_2, \ldots, \beta_k)'$ is $((k-1) \times 1)$, $\beta_1 = (\beta_{11}, \ldots, \beta_{1N})'$ is $(N \times 1)$, $X = [I_N \otimes i_T \quad X_s]$ is $(NT \times (k-1) + N)$ and $\beta = (\beta_1, \beta_s)'$ is $((K-1) + N \times 1)$. Note that $X_s = \begin{pmatrix} x_{21t}, x_{31t}, \ldots, x_{k1t} \\ x_{22t}, x_{32t}, \ldots, x_{k2t} \\ \vdots \; \vdots \; \vdots \; \vdots \\ x_{2nt}, x_{3nt}, \ldots, x_{knt} \end{pmatrix}$ and $var(e) = \Sigma$.
Given $X'X$ is non-singular and Σ a diagonal matrix with constant elements, the Ordinary Least Squares (OLS) estimator of β is

$$\hat{\beta} = (X'X)^{-1} X'y. \tag{32}$$

Computation of the Delete-One Jackknife estimator

1. Let $\hat{\beta}_{(it)}$ be the itth jackknife (OLS) estimate of β obtained by re-computing $\hat{\beta}$ in equation (32) with the itth group $(y_{it}, 1, x_{2it}, \ldots, x_{kit})$ deleted from the sample.

2. Compute the jackknife estimator as

$$\hat{\beta}_{(.)} = \frac{\Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \hat{\beta}_{(it)}}{NT}. \tag{33}$$

3. The jackknife estimate of the standard error is defined by

$$se_{jack}(\hat{\beta}_{(it)}) = \left[\frac{N-1}{NT} \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} (\hat{\beta}_{(it)} - \hat{\beta}_{(.)})^2 \right]^{\frac{1}{2}}. \tag{34}$$

Since $\hat{\beta}_{(it)}$ is a $((k-1) + N \times 1)$ multiparameter vector we can choose our parameter of interest as $\hat{\beta}_{(kit)}$, ie the kth element of $\hat{\beta}_{(it)}$ (see above note on confidence intervals for multiparameter vectors). Then a simple expression for the single parameter accelerator constant is

$$\hat{a} = \frac{\Sigma_{i=1}^{N} \Sigma_{t=1}^{T} (\hat{\beta}_{(k.)} - \hat{\beta}_{(kit)})^3}{6 \left[N \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} (\hat{\beta}_{(k.)} - \hat{\beta}_{(kit)})^2 \right]^{\frac{3}{2}}}. \tag{35}$$

where $\hat{\beta}_{(k.)}$ is the kth element of $\hat{\beta}_{(.)}$.
The confidence intervals mentioned so far, ie the standard or asymptotic,

percentile, BC and BC_a work well if $\hat{\beta}_k$ or some transformation of it has an approximate Gaussian distribution and the other parameters of the model satisfy some relatively simple assumptions. These are called transformation repecting properties of the interval and allow modifications and improvements to be made to the interval. See also Efron (1987).

(E) The Bootstrap-t Method

This was introduced in Efron (1982) and detailed in Efron and Tibshirani (1993), the bootstrap-t improves upon the percentile method. It is less computer intensive than the double bootstrap[2] and easier to impliment than the BC and BC_a methods, ie it involves no difficult computations. It is called the "Percentile-t" by Hall (1992). See Dicciccio and Romano (1988) for a review of bootstrap confidence intervals.
Consider the standard confidence interval derived in the previous section. Starting with $Z = \frac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \sim N(0,1)$. This led to the exact confidence interval

$$\left[\hat{\beta}_k - z^{(1-\alpha)} se(\hat{\beta}_k), \hat{\beta}_k - z^{(\alpha)} se(\hat{\beta}_k)\right]. \tag{36}$$

We know that when we use plug-in estimates, when the variance of $\hat{\beta}_k$ is unknown, that this interval holds asymptotically only in large samples. In finite samples, then, we obtain only approximate confidence intervals. For small samples the approximation of Z was improved upon by W. Gosset in 1908 with his Student's t-distribution. Now for small samples of size n with plug-in estimates for $var(\hat{\beta}_k)$ we have $Z_t = \frac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \sim t(n-1)$. Here $t(n-1)$ means Student's t-distribution with $(n-1)$ degrees of freedom (d.f.). Also $t(n-1) \rightarrow N(0,1)$ as $n \rightarrow \infty$. The percentiles of the t-distribution for varying degrees of freedom are tabulated in the Student's t-Tables.
Let $t^{(\alpha)}_{(n-1)}$ denote the αth percentile of the Student's t-distribution with $(n-1)$ d.f. Then our approximate $(1-2\alpha)$ confidence interval is

$$\left[\hat{\beta}_k - t^{(1-\alpha)}_{(n-1)} se(\hat{\beta}_k), \hat{\beta}_k - t^{(\alpha)}_{(n-1)} se(\hat{\beta}_k)\right]. \tag{37}$$

Our bootstrap-t interval is a generalisation of the above Student's t interval. The procedure estimates the distribution of Z_t directly from the data and tabulates percentiles that are appropriate for the data at hand.

The Bootstrap-t Algorithm

[2]The double bootstrap is sometimes called bootstrap iteration. It was developed in Hall and Martin (1988), Martin (1990) and Hall (1992) and is another way to improve on interval accuracy. It is a second-order accurate method. It is not shown here for reasons of brevity but is left to a future version of the paper

1. Generate B independent bootstrap data sets $X^{*1}, X^{*2}, \ldots, X^{*B}$, where $X^{*j} = \{X_{it}^{*j}, i = 1, \ldots, N, t = 1, \ldots, T\}$ and $j = 1, \ldots, B$.

2. Compute the bootstrap replication

$$Z^*(b) = \frac{\hat{\beta}^*(b) - \hat{\beta}_k}{\hat{se}^*(b)} \qquad \text{for} \quad b = 1, \ldots, B. \tag{38}$$

   Where as above $\hat{\beta}^*(b)$ is the value of $\hat{\beta}_k$ for the bootstrap sample X^{*b} and $\hat{se}^*(b)$ is the estimated standard error of $\hat{\beta}^*(b)$ for the bootstrap sample X^{*b}. N.B. $\hat{se}^*(b)$ is estimated as the regression $se(\hat{\beta}_k)$ from each bootstrap sample X^{*b} regression.

3. Let Z^* be the ordered list of $Z^*(b)$ replications. Estimate the αth percentile of the sorted vector of $Z^*(b)$'s by the value $\hat{t}^{(\alpha)}$ such that

$$\left(\# \frac{\{Z^*(b) \leq \hat{t}^{(\alpha)}\}}{B} \right) = \alpha. \tag{39}$$

   Thus if $B = 1000$ and $\alpha = 5\%$ then the 100αth empirical percentile is the $B.\alpha$th $= 1000.(0.05) = 50$th value of the ordered list of $Z^*(b)$ replications (or Z^*). Also for $\alpha = 95\%$, this gives the 950th value of the ordered list of $Z^*(b)$'s or (or Z^*). Thus the bootstrap-t confidence interval is given by

$$\left[\hat{\beta}_k - \hat{t}^{(1-\alpha)} se(\hat{\beta}_k), \hat{\beta}_k - \hat{t}^{(\alpha)} se(\hat{\beta}_k)\right] \tag{40}$$

   where $\hat{t}^{(\alpha)}$ is the αth percentile of the Z^* distribution.

We can also write equation (40) using endpoints as

$$\hat{\theta}_{lo}^t = \left[\hat{\beta}_k - \hat{t}^{(1-\alpha)} se(\hat{\beta}_k)\right] \quad = \text{Lower Bound.} \tag{41}$$

$$\hat{\theta}_{up}^t = \left[\hat{\beta}_k - \hat{t}^{(\alpha)} se(\hat{\beta}_k)\right] \quad = \text{Upper Bound.} \tag{42}$$

Hence the $(1 - 2\alpha)$ bootstrap-t confidence interval becomes $\left[\hat{\theta}_{lo}^t, \hat{\theta}_{up}^t\right]$.

4 The Dynamic OLS (DOLS) Estimation of a Panel Cointegration Regression

Although there has been great interest in testing for unit roots and cointegration in time-series data it is only recently that attention has been paid to

testing for unit roots and cointegration in panel data. Quah (1994), Levin and Lin (1992),(1993) and Im, Pesaran and Shin (1997) are a few studies of unit roots tests with panel data. Concerning panel cointegration tests we have the papers of McCoskey and Kao (1998), Kao (1999), Pedroni (1999) and Larsson, Lyhagen and Lothgren (2001). With the increasing use of non-stationary panel data the focus of panel data econometrics has shifted towards the study of the asymptotics of macro panels, with large N (eg individuals) and large T (eg time-series), as opposed to the usual asymptotics of micro panels with large N and small T. This has necessitated the development of a new limit theory for non-stationary panel data, ie limit distributions for double indexed integrated processes, by Phillips and Moon (1999a),(1999b). It was found that the statistical properties of the non-stationary panel data were very different from those of the non-stationary time-series data. The differences in the asymptotic statistical properties of the non-stationary panels have been highlighted by Kao and Chiang (1998),(2000), Phillips and Moon (1999a) and Pedroni (1996) in their works on the panel Fully Modified OLS (FMOLS), DOLS and OLS panel cointegration estimators. These works extend the field of panel cointegration to the estimation and inference of cointegrated regressions with panel data.
The FMOLS estimator of Phillips and Moon (1999a) and Pedroni (2000) is the panel analogue of the Phillips and Hansen (1990) FMOLS estimator of the time-series literature. These FMOLS estimators use non-parametric corrections for bias and endogeniety problems in the OLS estimator. Similarly the DOLS estimator of Kao and Chiang (1998),(2000) and Mark and Sul (1999), can be seen as the panel analogue of the Saikkonen (1991) and Stock and Watson (1993), DOLS estimator of the time-series literature. These DOLS estimators add leads and lags of the differenced regressors into the regression as parametric corrections for the bias and endogeniety problems. They are asymptotically equivalent to their FMOLS counterparts.
Since the introduction of these panel cointegation estimators a few Monte Carlo simulation studies of their finite sample properties, and some empirical applications, have appeared in the panel data literature. In a simulation study Kao and Chiang (2000) find i) OLS has a bias in finite samples ii) FMOLS does not improve on OLS in general iii) DOLS seems more promising than OLS or FMOLS in estimating panel cointegration regressions. Kao, Chiang and Chen (1999) apply the panel cointegration methods developed in Kao and Chiang (2000) to study R&D spillovers. They find FMOLS and DOLS produce slightly different results but are unanimous on the main issues. Funk (1998) also studies the same R&D spillovers using the panel cointegration methods developed by Kao (1999), Kao and Chiang (2000) and Pesaran, Shin and Smith (1999). Pedroni (2001) develops group mean

DOLS and FMOLS estimators which are the average of the individual time-series DOLS and FMOLS estimators. He compares his DOLS estimator with the ones of Kao and Chiang (2000) and of Mark and Sul (1999).

In this paper we use the pooled DOLS panel estimator of Kao and Chiang (2000) and the group-mean DOLS panel estimator of Pedroni (2001) to construct bootstrap confidence intervals for a purchasing power parity (PPP) panel cointegration regression. The question whether PPP holds in the long-run has been the focus of much empirical research, notably Frenkel (1981a), Frenkel (1981b), Messe and Rogoff (1988) and Officer (1982) in the time-series literature. Also recently panel methods have been used to test for PPP with great success. Here we have the papers by Oh (1996), Papell (1997), Wu (1996), McDonald (1996) and Coakley and Feurtes (1997). If long-run PPP is to hold (ie in the long-run the nominal exchange rate moves to equilibrate prices between countries), then there should be a long-run equilibrium (cointegrating) relation between the nominal exchange rate and the prices of domestic and foreign goods. This is testable using the following time-series regression

$$e_t = \alpha + \beta_1 p_t + \beta_2 p_t^* + u_t, \tag{43}$$

where $\{e_t, p_t$ and $p_t^*\}$ are the logarithms of the nominal exchange rate, domestic prices and foreign price level, respectively and $\{u_t\}$ is white noise. We proceed, using the two step method of Engle and Granger (1987), by first testing that each of e_t, p_t and p_t^* is I(1) and then show that some linear combination of them, ie a cointegrating regression, is I(0). If long run PPP holds then e_t should be cointegrated with p_t and p_t^*.

At the panel level most PPP tests have been conducted using panel unit roots tests (see the above panel PPP citations). However some use a panel version of the multivariate VECM framework proposed by Johansen (1988),(1991),(1995). The results are mixed but generally in favour of PPP using a strong form of the hypothesis which involves the joint symmetry and proportionality assumption $\beta_1 = -\beta_2 = 1$. Some studies support a weak form of PPP where the β_i coefficients fall within the range $(0 < \beta_1 < 2, -2 < \beta_2 < 0)$. Some PPP Panel VECM studies have come from Jacobson, Lyhagen, Larsson and Nessen (2002) and Banerjee, Marcellino and Osbat (2001). Jacobson, Lyhagen, Larsson and Nessen (2002) found some support for a weak form of PPP in their panel of four OECD countries. Finally Pedroni (2001) investigates the PPP hypothesis for a panel of 20 developed countries and the FMOLS estimators discussed above. He finds strong support against a strong form of the PPP hypothesis.

Given the two main PPP hypotheses it is interesting to highlight the theories in a confidence interval framework. Thus given point estimates close to unity and a small variability of the interval estimates would lend support to the strong PPP hypothesis. Whereas point estimates falling within the range $(0 < \beta_1 < 2, -2 < \beta_2 < 0)$ and interval estimates with large variability would lend support to the weak PPP hypothesis. One of the panel cointegration regressions is constructed by pooling the cross-section dimension and assuming heterogeniety in the intercepts, see Kao (1999), Kao, Chiang and Chen (1999) and Kao and Chiang (2000). Thus by assuming homogeniety of β across panels we in effect impose a strong PPP hypothesis, if it turns out that $\beta_1 = -\beta_2 = 1$ for all units. The other panel cointegration regression of Pedroni (2000), (2001) differs in that he uses heterogeneous panels where the β coefficients are allowed to vary across individuals or countries.

5 The DOLS Panel Regression Model and the Bootstrap

The Kao and Chiang (2000) DOLS Panel Estimator

Consider the fixed effects panel regression

(44) $$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$,
where $\{y_{it}\} \sim I(1)$ are (1×1) scalars, β is a $(k \times 1)$ parameter vector, α_i a (1×1) scalar of individual intercepts, $\{u_{it}\}$ a (1×1) stationary disturbance term, $\{x_{it}\}$ is a $(k \times 1)$ vector of integrated processes of order 1, $\forall i$ such that $x_{it} = x_{it-1} + \epsilon_{it}$, thus $x_{it} \sim I(1)$. Also $x_{i0} = y_{i0} = 0$ where independence is assumed of the $\{y_{it}\}$, $\{x_{it}\}$ and $\{u_{it}\}$ across i. Thus equation (44) describes a system of cointegrated regressions where $\{y_{it}\}$ is cointegrated with $\{x_{it}\}$. To estimate the equation by DOLS techniques observe the following. We write $\{u_{it}\}$, following Saikkonen (1991) as

(45) $$u_{it} = \Sigma_{j=-\infty}^{\infty} c_{ij}\epsilon_{it+j} + v_{it}$$

where v_{it} is a stationary random variable with zero mean and ϵ_{it} is an i.i.d. zero mean random variable. ϵ_{it} and v_{it} are uncorrelated contemporaneously and also in all leads and lags. In practice the leads and lags are truncated so that

(46) $$u_{it} = \Sigma_{j=-q}^{q} c_{ij}\epsilon_{it+j} + \bar{v}_{it}.$$

For a more detailed discussion see Kao and Chiang (2000). Given our large sample size, eg 2622 we expect that $\bar{v}_{it}$ is a good approximation of v_{it}. Substituting u_{it} in the above we get

(47) $$y_{it} = \alpha_i + x'_{it}\beta + \Sigma_{j=-q}^{q} c_{ij}\epsilon_{it+j} + \bar{v}_{it}$$

where $\bar{v}_{it} = v_{it} + \Sigma_{j=-q}^{q} c_{ij}\epsilon_{it+j}$. Also note $\Delta x_{it} = x_{it} - x_{it-1}$. Thus to obtain the DOLS estimate of β we run the following regression

$$y_{it} = \alpha_i + x'_{it}\beta + \Sigma_{j=-q}^{q} c_{ij}\Delta x_{it+j} + \bar{v}_{it}. \tag{48}$$

Kao and Chiang (2000) show that the DOLS estimator has the same limiting distribution as FMOLS. Moreover an important departure from the existing time-series literature is that the asymptotics are calculated using the sequential limit theorems of Phillips and Moon (1999a). Another point worth mentioning is that in Kao, Chiang and Chen (1999) it is noted that there is not yet a coherent strategy to be applied for estimating the lengths of lags and leads in these panel cointegration models. The method used in this paper was the general-to-specific method advocated by D.F. Hendry. Here we start with an overparameterised model and use sequential test procedures to test down for a more parsimonious representation. In practice this meant setting leads and lags of 3 for both regressors and then testing for their significance. All insignificant regressors were subsequently dropped from the regression. For estimation purposes we have the panel analogue of equation (43)

$$e_{it} = \alpha_i + \beta_1 p_{it} + \beta_2 p^*_{it} + u_{it}, \tag{49}$$

where $\{e_{it}, p_{it}$ and $p^*_{it}\} \sim I(1)$ are the logarithms of the nominal exchange rate, domestic prices and foreign price level, respectively for country i at time t and similarly u_{it} is a stationary disturbance term. Given the DOLS transformation we get the final regression of

$$e_{it} = \alpha_i + \beta_1 p_{it} + \beta_2 p^*_{it} + \Sigma_{j=-q}^{q} d_{1ij}\Delta p_{it+j} + \Sigma_{j=-q}^{q} d_{2ij}\Delta p^*_{it+j} + u_{it}. \tag{50}$$

The Pedroni DOLS Panel Estimator

In his paper Pedroni (2001) develops group-mean FMOLS and DOLS panel estimators for cointegration vectors in heterogeneous panels. These estimators are different to the Kao and Chiang (2000) and Mark and Sul (1999) panel estimaors, which he calls within-dimension estimators, in that his estimators are constructed using the between-dimension of the panel. This method allows for greater flexibility when estimating cointegrating vectors in that hetrogeniety is allowed amongst the individual members of the panel. Both the Pedroni (2001) group-mean FMOLS and DOLS panel estimators are formed by averaging over the individual FMOLS and DOLS time-series estimators applied to the ith member of the panel. We now show how the panel DOLS estimator is formed. Using regression (50) Pedroni constructs his group-mean DOLS panel estimator as follows

$$\hat{\beta}^{*}_{GD} = \left[N^{-1}\Sigma_{i=1}^{N} \left(\Sigma_{t=1}^{T} z_{it} z'_{it} \right)^{-1} \left(\Sigma_{t=1}^{T} z_{it} \tilde{e}_{it} \right) \right]_1 , \tag{51}$$

where z_{it} is the $(4(K+1) \times 1)$ vector of regressors

$$z_{it} = ((p_{it} - \bar{p}_i), (p^{*}_{it} - \bar{p}^{*}_i), \triangle p_{it-K}, \ldots, \triangle p_{it+K}, \triangle p^{*}_{it-K}, \ldots, \triangle p^{*}_{it+K}) \tag{52}$$

and $\tilde{e}_{it} = e_{it} - \bar{e}_i$. Here $\bar{e}_i = \frac{\Sigma_{t=1}^{T} e_{it}}{T}$ and similarly $\bar{p}_i = \frac{\Sigma_{t=1}^{T} p_{it}}{T}$. The subscript 1 outside the brackets indicate that we are considering only the first element of the vector for the pooled slope coefficient. The estimator can also be written simply as

$$\hat{\beta}^{*}_{GD} = N^{-1}\Sigma_{i=1}^{N} \hat{\beta}^{*}_{Di} \tag{53}$$

where $\hat{\beta}^{*}_{Di}$ is the conventional DOLS time-series estimator applied to the ith member of the panel.

Let $\sigma_i^2 = \lim_{T \longrightarrow \infty} E\left[T^{-1} (\Sigma_{t=1}^{T} \hat{u}_{it})^2 \right]$ be the long-run variance of the residuals from the DOLS regression. This can be estimated using standard HAC methods, such as the Newey-West HAC estimator. Then the t-statistic for the between-dimension estimator is written

$$t_{\hat{\beta}^{*}_{GD}} = N^{-0.5}\Sigma_{i=1}^{N} t_{\hat{\beta}^{*}_{Di}} \tag{54}$$

where[3]

$$t_{\hat{\beta}^{*}_{Di}} = (\hat{\beta}^{*}_{Di} - \beta) \left(\hat{\sigma}_i^{-2} \Sigma_{t=1}^{T} (\tilde{p}_{it} - \bar{\tilde{p}})(\tilde{p}_{it} - \bar{\tilde{p}})' \right)^{0.5} , \tag{55}$$

where $(\tilde{p}_{it} - \bar{\tilde{p}}_i) = (p_{it} - \bar{p}_i, p^{*}_{it} - \bar{p}^{*}_i)$.

The Data Set

The data set[4] is quarterly observations over the period 1957Q1-1991Q2, for 20 OECD countries, taken from the IMF International Financial Statistics.[5] The variables used were, p_t the consumer price level (or CPI), e_t was taken as the market rate per U.S. Dollar and p^{*}_t the U.S. consumer price level (or CPI).

[3] In the computations the simple average of the Pedroni t-statistics was used

[4] The OECD countries included in the panel are Austria, Belgium, Canada, Switzerland, Germany, Spain, France, U.K., Greece, Ireland, Italy, Japan, Luxembourg, Norway, Sweden, U.S., Portugal, Iceland and Denmark. In the panel cointegration regression the U.S. is used as numeraire.

[5] Obtained from the MIMAS archives

The Bootstrap Method for Cointegrating Regressions

The simple bootstrap method of Efron (1979) was originally designed for i.i.d. errors. When using time-series models, such as unit root and cointegration models, the bootstrap methodology needs to be modified to cope with errors that not be i.i.d., eg the stationarity assumption of u_{it} might range from white noise or weak stationarity to an m-dependent, α or strong mixing sequence. Li and Maddala (1996),(1997) discuss a number of bootstrap methods that are applicable to time-series models. In particular in Li and Maddala (1997) cointegrating regressions are studied and the appropriate bootstrap method considered. Discussed are the recursive bootstrap, the moving blocks (MBB) bootstrap and the stationary (SB) bootstrap and it is explained which method suits a particular situation. Li and Maddala (1996),(1997) also discuss the choice of procedure for the generation of the bootstrap samples when using cointegrating regressions and highlight the choice between the direct method of bootstrapping the data or the alternative of bootstrapping the residuals. They explain that for cointegrating regressions only the latter is appropriate. The basic argument is that all the information of the structure of the model should be used when generating the bootstrap samples. Only when the residual bootstrap method is used is this condition satisfied. Li and Maddala suggest the pairs bootstrap method for bootstrapping the residuals in the cointegrating regression. Thus estimate equation (50) by DOLS and obtain residuals $\hat{u}_{it}$. Noting that $\hat{\beta}_1$ and $\hat{\beta}_2$ are super consistent obtain also the residuals $\hat{w}_{1it} = \triangle p_{it}$ and $\hat{w}_{2it} = \triangle p^*_{it}$. After centering the residuals bootstrap the pairs $(\hat{u}_{it}, \hat{w}_{it})$ where $\hat{w}_{it} = (\hat{w}_{1it}, \hat{w}_{2it})$. Next construct the bootstrap samples of p^*_{it} and p^{**}_{it} recursively and finally using $\hat{\beta}_1, \hat{\beta}_2, u^*_{it}, p^*_{it}$ and $p^{**}_{it}, etc.\forall i, t$, construct the sample $e^* = (e^*_{11}, e^*_{12}, \ldots, e^*_{it})^{/}$ for use in equation (50). Finally Li and Maddala (1997) conduct a Monte Carlo experiment to compare the asymptotic FMOLS methods with MBB and SB methods using FMOLS, in a cointegrating regression with serial correlation in the errors and endogeniety of the regressors. They conclude that the residual bootstrap performs poorly if the serial correlation structure of the errors is mispecified. Despite the lack of a theoretical basis for the MBB and SB (there are no theoretical results proved yet for the MBB in cointegrating regressions) in simulations these methods seem to improve significantly on asymptotic inference and are thus recommended for use by empirical researchers. Chang, Park and Song (2002) also study in detail the topic of bootstrapping cointegrating regressions. They employ the sieve bootstrap method coupled with the pairs bootstrap for generating bootstrap samples. They also give a rigourous exposition of the asymptotic theory of the sieve bootstrap approach to bootstrapping cointegrating regressions. Both the OLS and DOLS estimator of

Saikkonen (1991) are discussed. Finally a Monte Carlo study to investigate the finite sample performance of the bootstrap methods considered is carried out.
In this paper we explore some of the issues raised by Li and Maddala (1996) and (1997). We compare three different approaches to bootstrapping a panel cointegration regression. The first is based on the simple bootstrap and the Kao and Chiang (2000) panel DOLS estimator. Although disallowed by Li and Maddala (1997) this method is justified here by the use of a very large sample size. The panel contains over 2600 observations and Li and Maddala's (1997) discussion concentrated on small sample sizes. This method has been used successfully by Shea (1989) and more recently by Kilian (1999). In the latter the very large sample size (of 2000 observations) was crucial in the ability to weaken the i.i.d. error term assumption usually required for the simple bootstrap. The second is based on the Pedroni (2001) group-mean panel DOLS estimator discussed above. Here we apply the pairs bootstrap to each individual country in the panel and use the bootstrap methodology to compute group averages. Finally the pooled panel DOLS estimator of Kao and Chiang (2000) is studied using the pairs bootstrap method.
As explained by Li and Maddala (1997) earlier, complications arise if there is serial correlation in the residuals of the cointegration regression. Hence in the presence of serial correlation of the errors the pairs bootstrap should be modified to take this into account. This is exactly what Li and Maddala (1997) suggest in their paper. If the auto-correlation structures of $\hat{u}_{it}$ are known then a recursive bootstrap can be applied to them, in addition to the pairs bootstrap. Otherwise for general unknown serial correlation one can use the moving block bootstrap (MBB), in addition to the pairs bootstrap. In our estimation results we detected severe serial correlation in the residuals of our DOLS regression estimates. The serial correlation was later specified to be of AR(1) form. Hence a recursive bootstrap method was used for the AR(1) residuals of the cointegration regression. This modification of the pairs bootstrap is included in our bootstrap algorithms given below.

The Bootstrap Algorithms[6]

The Simple Bootstrap

The simple bootstrap procedure is carried out by resampling the errors from the estimated equation (48) or (50).

1. Compute the predicted residuals using the estimates from equation

[6] All bootstrap computations were carried out using the Ox programming language

(48), ie $\hat{\alpha}_i, \hat{\beta}$ and $\hat{c}_{ij}$ thus

$$\hat{\hat{v}}_{it} = \hat{\alpha}_i + x_{it}'\hat{\beta} + \Sigma_{j=-q}^{q}\hat{c}_{ij}\Delta x_{it+j} \tag{56}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

2. Obtain the vector of residuals $\hat{\hat{v}} = (\hat{\hat{v}}_{i1}, \hat{\hat{v}}_{i2}, \ldots, \hat{\hat{v}}_{NT})'$ and recentre the residuals using

$$\hat{\hat{v}}_{(.)} = \frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{i=1}^{N}\hat{\hat{v}}_{it}. \tag{57}$$

Thus the vector of recentred residuals is $\hat{\hat{v}}^c = (\hat{\hat{v}}_{i1}^c, \hat{\hat{v}}_{i2}^c, \ldots, \hat{\hat{v}}_{NT}^c)'$ where

$$\hat{\hat{v}}_{i1}^c = (\hat{\hat{v}}_i - \hat{\hat{v}}_{(.)}), \hat{\hat{v}}_{i2}^c = (\hat{\hat{v}}_{i2} - \hat{\hat{v}}_{(.)}), \ldots, \hat{\hat{v}}_{NT}^c = (\hat{\hat{v}}_{NT} - \hat{\hat{v}}_{(.)}). \tag{58}$$

3. Resample from the vector of centred residuals with replacement. So that v^* is the vector of resampled residuals

$$v^* = (v_{i1}^*, v_{i2}^*, \ldots, v_{NT}^*)'. \tag{59}$$

4. Construct the bootstrap samples from

$$y_{it}^* = \hat{\alpha}_i + x_{it}'\hat{\beta} + \Sigma_{j=-q}^{q}\hat{c}_{ij}\Delta x_{it+j} + v_{it}^*. \tag{60}$$

5. Using the bootstrap samples y_{it}^*, α_i and x_{it} estimate $\beta^*(b)$ the bootstrap DOLS estimate of β.

6. Repeat steps (2) to (5) B times.

7. Compute the bootstrap point estimator

$$\beta_B^* = \frac{\Sigma_{b=1}^{B}\beta^*(b)}{B}. \tag{61}$$

8. Compute the bootstrap variance of β_B^*

$$var\,(\beta_B^*) = \frac{\Sigma_{b=1}^{B}(\beta^*(b) - \beta_B^*)^2}{B-1}. \tag{62}$$

NB In the B repetitions the variance of $\beta^*(b)$ is taken to be the regression estimate $var(\hat{\beta})$ since $var(\beta^*(b)) \to var(\hat{\beta})$ as $B \to \infty$.

The Pairs Bootstrap and the Pooled DOLS estimator

The pairs bootstrap procedure is also carried out by resampling the errors from the estimated equation (48) or (50) and the stochastic error terms of the I(1) regressors.

1. Again compute the predicted residuals using the estimates from equation (48), ie $\hat{\alpha}_i, \hat{\beta}$ and $\hat{c}_{ij}$ thus

   (63) $$\hat{\hat{v}}_{it} = \hat{\alpha}_i + x'_{it}\hat{\beta} + \Sigma_{j=-q}^{q}\hat{c}_{ij}\Delta x_{it+j}$$

   for $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

2. Obtain the residuals $\hat{w}_{1it} = \triangle x_{1it}, \ldots, \hat{w}_{kit} = \triangle x_{kit}$ and form the vector $\hat{w}_{it} = (\hat{w}_{1it}, \ldots, \hat{w}_{kit})$ and $\hat{w} = (\hat{w}_{i1}, \hat{w}_{i2}, \ldots, \hat{w}_{NT})'$. Recentre these using

   (64) $$\hat{w}_{(.)} = \frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{i=1}^{N}\hat{w}_{it}$$

   for $i = 1, \ldots, N$ and $t = 1, \ldots, T$.
   Thus we have $\hat{w}^c = (\hat{w}_{i1}^c, \hat{w}_{i2}^c, \ldots, \hat{w}_{NT}^c)'$ where

   (65) $$\hat{w}_{i1}^c = (\hat{w}_{i1} - \hat{w}_{(.)}), \hat{w}_{i2}^c = (\hat{w}_{i2} - \hat{w}_{(.)}), \ldots, \hat{w}_{NT}^c = (\hat{w}_{NT} - \hat{w}_{(.)}).$$

3. If we assume that the residuals $\hat{\hat{v}}_{it}$ follow an AR(1) process then

   (66) $$\hat{\hat{v}}_{it} = \rho\hat{\hat{v}}_{it-1} + \varepsilon_{it}$$

   where ε_{it} is a white noise error term.
   Run the regreassion of equation (66) and obtain the residuals $\hat{\varepsilon}_{it}$ and also $\hat{\rho}$ and the vector of residuals $\hat{\varepsilon} = (\hat{\varepsilon}_{i1}, \hat{\varepsilon}_{i2}, \ldots, \hat{\varepsilon}_{NT})'$. Recentre these using

   (67) $$\hat{\varepsilon}_{(.)} = \frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{i=1}^{N}\hat{\varepsilon}_{it}.$$

   Thus the vector of recentred residuals is $\hat{\varepsilon}^c = (\hat{\varepsilon}_{i1}^c, \hat{\varepsilon}_{i2}^c, \ldots, \hat{\varepsilon}_{NT}^c)'$ where

   (68) $$\hat{\varepsilon}_{i1}^c = (\hat{\varepsilon}_{i1} - \hat{\varepsilon}_{(.)}), \hat{\varepsilon}_{i2}^c = (\hat{\varepsilon}_{i2} - \hat{\varepsilon}_{(.)}), \ldots, \hat{\varepsilon}_{NT}^c = (\hat{\varepsilon}_{NT} - \hat{\varepsilon}_{(.)}).$$

4. Resample with replacement from the paired vector of recentred residuals, $\hat{z}_{it}^c = (\hat{\varepsilon}_{it}^c, \hat{w}_{it}^c)$ to get the bootstrap sample $z_{it}^* = (\varepsilon_{it}^*, w_{it}^*)$. So that ε^* and w^* are the vectors of resampled residuals

   (69) $$\varepsilon^* = (\varepsilon_{i1}^*, \varepsilon_{i2}^*, \ldots, \varepsilon_{NT}^*)' \quad \text{and} \quad w^* = (w_{i1}^*, w_{i2}^*, \ldots, w_{NT}^*)'.$$

5. Obtain the bootstrap samples of $x^*_{1it}, \dots, x^*_{kit}$ by recursion using the initial conditions $x^*_{k00} = x_{k00}$. That is

$$x^*_{1it} = x^*_{1it-1} + w^*_{1it}, \dots, x^*_{kit} = x^*_{kit-1} + w^*_{kit}. \tag{70}$$

Alternatively form x^*_{kit} from

$$x^*_{kit} = x^*_{k00} + \Sigma^T_{t=1}\Sigma^N_{i=1} w^*_{kit}. \tag{71}$$

Also obtain the bootstrap samples of v^*_{it} by recursion using the estimated $\hat{\rho}$ of equation (66) and the initial conditions $v^*_{00} = \hat{\hat{v}}_{00}$. Thus

$$v^*_{it} = \hat{\rho} v^*_{it-1} + \varepsilon^*_{it}. \tag{72}$$

6. Construct the bootstrap samples of y^*_{it} from

$$y^*_{it} = \hat{\alpha}_i + x^{*\prime}_{it}\hat{\beta} + \Sigma^q_{j=-q}\hat{c}_{ij}\Delta x^*_{it+j} + v^*_{it}. \tag{73}$$

7. Using the bootstrap samples y^*_{it}, α_i and x^*_{it} estimate $\beta^*(b)$ the bootstrap DOLS estimate of β.

8. Repeat steps (2) to (7) B times.

9. Compute the bootstrap point estimator

$$\beta^*_B = \frac{\Sigma^B_{b=1}\beta^*(b)}{B}. \tag{74}$$

10. Compute the bootstrap variance of β^*_B

$$var\,(\beta^*_B) = \frac{\Sigma^B_{b=1}(\beta^*(b) - \beta^*_B)^2}{B-1}. \tag{75}$$

NB Again in the B repetitions the variance of $\beta^*(b)$ is taken to be the regression estimate $var(\hat{\beta})$.

The Pairs Bootstrap and the Group-Mean DOLS estimator

The pairs bootstrap procedure for the group-mean DOLS estimator is carried out by resampling the errors from the individual country estimates and I(1) regressor stochastic error terms and averaging over the panel.

1. Using equation (48) compute the DOLS estimates for the ith individual country and obtain the predicted residuals using $\hat{\alpha}, \hat{\beta}$ and $\hat{c}_j$ thus

$$\hat{\hat{v}}_t = \hat{\alpha} + x'_t\hat{\beta} + \Sigma^q_{j=-q}\hat{c}_j\Delta x_{t+j} \tag{76}$$

for $t = 1, \dots, T$.

2. Obtain the residuals $\hat{w}_{1t} = \triangle x_{1t}, \ldots, \hat{w}_{kt} = \triangle x_{kt}$ and form the vector $\hat{w}_t = (\hat{w}_{1t}, \ldots, \hat{w}_{kt})$ and $\hat{w} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_T)'$. Recentre these using

$$\hat{w}_{(.)} = \frac{1}{T}\Sigma_{t=1}^{T}\hat{w}_t \tag{77}$$

for $t = 1, \ldots, T$.
Thus we have $\hat{w}^c = (\hat{w}_1^c, \hat{w}_2^c, \ldots, \hat{w}_T^c)'$ where

$$\hat{w}_1^c = (\hat{w}_1 - \hat{w}_{(.)}), \hat{w}_2^c = (\hat{w}_2 - \hat{w}_{(.)}), \ldots, \hat{w}_T^c = (\hat{w}_T - \hat{w}_{(.)}). \tag{78}$$

3. Similarly if we again assume that the residuals $\hat{\hat{v}}_t$ follow an AR(1) process then

$$\hat{\hat{v}}_t = \rho\hat{\hat{v}}_{t-1} + \varepsilon_t \tag{79}$$

where again ε_t is a white noise error term.
Run the regreassion of equation (79) and obtain the residuals $\hat{\varepsilon}_t$ and also $\hat{\rho}$ and the vector of residuals $\hat{\varepsilon} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_T)'$. Recentre these using

$$\hat{\varepsilon}_{(.)} = \frac{1}{T}\Sigma_{t=1}^{T}\hat{\varepsilon}_t. \tag{80}$$

Thus the vector of recentred residuals is $\hat{\varepsilon}^c = (\hat{\varepsilon}_1^c, \hat{\varepsilon}_2^c, \ldots, \hat{\varepsilon}_T^c)'$ where

$$\hat{\varepsilon}_1^c = (\hat{\varepsilon}_1 - \hat{\varepsilon}_{(.)}), \hat{\varepsilon}_2^c = (\hat{\varepsilon}_2 - \hat{\varepsilon}_{(.)}), \ldots, \hat{\varepsilon}_T^c = (\hat{\varepsilon}_T - \hat{\varepsilon}_{(.)}). \tag{81}$$

4. Resample with replacement from the paired vector of recentred residuals, $\hat{z}_t^c = (\hat{\varepsilon}_t^c, \hat{w}_t^c)$ to get the bootstrap sample $z_t^* = (\varepsilon_t^*, w_t^*)$. So that ε^* and w^* are the vectors of resampled residuals

$$\varepsilon^* = (\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_T^*)' \quad \text{and} \quad w^* = (w_1^*, w_2^*, \ldots, w_T^*)'. \tag{82}$$

5. Obtain the bootstrap samples of $x_{1t}^*, \ldots, x_{kt}^*$ by recursion using the initial conditions $x_{k0}^* = x_{k0}$. That is

$$x_{1t}^* = x_{1t-1}^* + w_{1t}^*, \ldots, x_{kt}^* = x_{kt-1}^* + w_{kt}^*. \tag{83}$$

Alternatively form x_{kt}^* from

$$x_{kt}^* = x_{k0}^* + \Sigma_{t=1}^{T} w_{kt}^*. \tag{84}$$

Also obtain the bootstrap samples of v_t^* by recursion using the estimated $\hat{\rho}$ of equation (79) and the initial conditions $v_0^* = \hat{\hat{v}}_0$. Thus

$$v_t^* = \hat{\rho}v_{t-1}^* + \varepsilon_t^*. \tag{85}$$

6. Construct the bootstrap samples of y_t^* from

$$y_t^* = \hat{\alpha} + x_t^{*\prime}\hat{\beta} + \Sigma_{j=-q}^{q}\hat{c}_j \Delta x_{t+j}^* + v_t^*. \tag{86}$$

7. Using the bootstrap samples y_t^*, α and x_t^* estimate $\beta^*(z)$ the bootstrap DOLS estimate of β for country z.

8. Repeat steps (2) to (7) for each country in the panel.

9. Compute the bootstrap group-mean β DOLS estimator as

$$\beta^*(b) = \frac{\Sigma_{z=1}^{N}\beta^*(z)}{N}. \tag{87}$$

10. Compute the bootstrap group-mean β DOLS t-statistic as[7]

$$t_{(\beta^*(b))} = \frac{\Sigma_{z=1}^{N} t^*(z)}{N}. \tag{88}$$

    where $t^*(z)$ is the bootstrap DOLS estimate of the t-statistic of $\beta^*(z)$ for country z.

11. For each panel of N countries construct B bootstrap samples using the initial regressions described above in step (1). That is repeat steps (2) to (10) B times.

12. Compute the bootstrap point estimator

$$\beta_B^* = \frac{\Sigma_{b=1}^{B}\beta^*(b)}{B}. \tag{89}$$

13. Compute the bootstrap variance of β_B^*

$$var\,(\beta_B^*) = \frac{\Sigma_{b=1}^{B}(\beta^*(b) - \beta_B^*)^2}{B-1}. \tag{90}$$

[7] For the group-mean β DOLS bootstrap-t confidence interval the t-statistic $Z^*(b)$ of equation (38) was used in the above procedure

6 Efficient Estimation

One of the purposes of this study is to see how well the bootstrap[8] works with non-stationary panel data. One would like to judge the performance and efficiency of the bootstrap in a general setting when compared with its counterparts from first-order asymptotic theory. We have already mentioned about the power of the bootstrap to deliver asymptotic refinements, these lead to more efficient estimates. There are methods of evaluating confidence intervals in order to make efficiency judgements. Two criterion are used to judge confidence intervals (i) size and (ii) coverage probability. An optimal and hence efficient confidence interval is one with small size and large coverage, but these are difficult to obtain. We may measure coverage probability by the true average coverage probability and size by the length of the interval. Efron (1990) considers an optimal confidence interval as one that is the shortest possible for a given coverage and calls it a, "correct" interval.
Consider again Corollory's (2.7) and (2.10) From the first we find that the difference between the true coverage probability and the nominal coverage probability of the asymptotic confidence interval is $O(n^{-1})$ that is

$$P(|T_n| \leq z_{(\infty,\frac{\alpha}{2})}) = 1 - \alpha + O(n^{-1}). \tag{91}$$

Whilst from the second we have that for the bootstrap this error is $O(n^{-2})$ given by

$$P(|T_n| \leq z^*_{(n,\frac{\alpha}{2})}) = 1 - \alpha + O(n^{-2}). \tag{92}$$

Note these confidence intervals are two-sided intervals. In general for one-sided and equal-tailed confidence intervals the difference between the true coverage probability and the nominal coverage probability for the asymptotic interval is $O(n^{-\frac{1}{2}})$. In this case the asmptotic confidence interval is said to be first-order accurate. However the analogous difference for the bootstrap confidence interval is $O(n^{-1})$. In which case we say it is second-order accurate. These notions of accuracy give us a good guide as to the expected performance of our confidence interval methodologies. We see that the errors are much smaller with the bootstrap than with the asymptotic confidence interval. The standard or asymptotic confidence interval and percentile intervals are first-order accurate. Whilst the BC_a and bootstrap-t

[8]Horowitz (2000) notes that efficient estimation often results when one uses the parametric bootstrap, as opposed to the non-parametric bootstrap. If one knows the parametric distribution of the errors being sampled then this, when used, will be more accurate than using the empirical distribution function and leads to smaller errors. With the parametric bootstrap an assumption is made as to the form of the distribution being sampled (eg normal). Whilst with the non-parametric no assumptions are made

intervals are second-order accurate. See Efron (1987) for a good discussion. The asymptotic theory for these accuracy, coverage probability and interval size, concepts has been rigourously proved using Edgeworth and Cornish Fisher expansions by Hall (1992).

7 Estimation Results

In Table 1 we have the DOLS estimation results for the pooled panel cointegration model. The high $R^2 = 0.99675$ indicates that the model is a very good fit of the data. However the regression D.W. statistic (not shown) is very low $D.W. = 0.108$ and so we should use the appropriate panel data tests for serial correlation.
To test for first-order serial correlation in a fixed effects model we use the LM test developped in Baltagi and Li (1995). Consider the model of equation (50) with the errors described by the AR(1) process

$$\hat{u}_{it} = \rho \hat{u}_{it-1} + \varepsilon_{it}, \tag{93}$$

where ε_{it} is a Gaussian white noise process. Under the null hypothesis, $H_0 : \rho = 0$ and $H_1 : \rho \neq 0$

$$LM_1 = \left[\frac{NT^2}{(T-1)} \right] \left(\frac{\Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \hat{u}_{it} \hat{u}_{it-1}}{\Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \hat{u}_{it}^2} \right) \tag{94}$$

As $T \longrightarrow \infty$ the statistic LM_1 is distributed as a χ_1^2 variable.
Our computed statistic $LM_1 = 2363.2$. With critical χ_1^2 values of 3.84 and 6.63 at the 5% and 1% significance levels, respectively, the null hypothesis is decisively rejected in favour of the hypothesis of serial correlation in the errors.
In order to test the assumption of homogeniety of the slope coefficients of the panel across individuals the Chow (1960) test for poolability of the data was carried out. Due to serial correlation we use the Chow test after GLS transformation. The restricted model is the pooled model where the slope coefficient is the same for all individuals. Whilst the unrestricted model has slope coefficients varying over individuals and amounts to a seperate regression for each country. We transform both the models using the Cochrane-Orcutt method applied to each individual country in the panel.
Let $u^* = (\hat{u}_{i1}^*, \hat{u}_{i2}^*, \ldots \hat{u}_{NT}^*)'$ be the vector of residuals of the transformed resticted model[9] and let $u_1^* = (\hat{u}_{11}^*, \hat{u}_{12}^*, \ldots \hat{u}_{1T}^*)'$, $u_2^* = (\hat{u}_{21}^*, \hat{u}_{22}^*, \ldots \hat{u}_{2T}^*)'$, and

[9] The estimation results for these transformed models are not shown

so on, be the vectors of residuals of the transformed unrestricted model. Given β is the vector of slope coefficients for the restricted model and β_i, for all i, those for the unrestricted model, respectively, then under the null hypothesis $H_0 : \beta_i = \beta$ and $H_1 : \beta_i \neq \beta \; \forall i$

$$\mathrm{F}_1 = \frac{(u^{*\prime}u^{*} - u_1^{*\prime}u_1^{*} - u_2^{*\prime}u_2^{*}, - \ldots - u_N^{*\prime}u_N^{*})}{(N-1)K'} \Big/ \frac{(u_1^{*\prime}u_1^{*} + u_2^{*\prime}u_2^{*}, + \ldots + u_N^{*\prime}u_N^{*})}{N(T-K')}$$

This F_1 statistic has an F-distribution with $(N-1)K'$ and $N(T-K')$ degrees of freedom, ie $F_1 \sim F((N-1)K', N(T-K'))$ and where $K' = K+1$. Here our computed $F_1 = 2.81623$ with critical $F(324, 2128)$ values of 1.11 and 1.15 at the at the 5% and 1% significance levels respectively. Hence the null hypothesis of a common slope coefficient across the panel is rejected.
In Tables 1 and 2 we show the Heteroscedastity and Autocorrelation Consistent (HAC) standard error estimates using the long-run variance of the error term. The first is the HAC1 Newey and West (1987) estimator outlined in Pedroni (2001) above as

$$\hat{V} = \left(\frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}X_{it}X_{it}'\right)^{-1} \times$$
$$\frac{1}{NT}\left(\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}\hat{u}_{it}^2 + \Sigma_{v=1}^{q}\left[1 - \frac{v}{q+1}\right]\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}\left(\hat{u}_{it}\hat{u}_{it-v} + \hat{u}_{it-v}\hat{u}_{it}\right)\right).$$

The HAC2 estimator is also a Newey and West (1987) estimator calculated as

$$\hat{V} = \left(\frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}X_{it}X_{it}'\right)^{-1} \times$$
$$\frac{1}{NT}\left(\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}\hat{u}_{it}^2X_{it}X_{it}' + \Sigma_{v=1}^{q}\left[1 - \frac{v}{q+1}\right]\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}\left(X_{it}\hat{u}_{it}\hat{u}_{it-v}X_{it-v}' + X_{it-v}\hat{u}_{it-v}\hat{u}_{it}X_{it}'\right)\right)$$
$$\times \left(\frac{1}{NT}\Sigma_{t=1}^{T}\Sigma_{N=1}^{N}X_{it}X_{it}'\right)^{-1}$$
where

$$(95)\; X_{it} = (i_1, i_2, \ldots i_{19}, p_{it}, p_{it}^{*}, \triangle p_{it-K}, \ldots, \triangle p_{it+K}, \triangle p_{it-K}^{*}, \ldots, \triangle p_{it+K}^{*})$$

The Barlett window was chosen to describe the lag structure of the Newey and West estimators with a truncation point of 10 (ie q=10). This was chosen after inspection of the sample autocorrelation function (ACF) of the residuals $\hat{u}_{it}$, shown in Figure (2). The graph depicts lagged correlations persisting even after 10 lags.
Individual ADF tests were carried out for each country on $\{e_t, p_t$ and $p_t^{*}\}$ to ensure they were $I(1)$. Often it is the case that nominal prices are I(2) necessitating second differencing to form an I(0) variable. The results are shown in Table 10. We see that for most countries p_t, p_t^{*}, and e_t can be safely termed I(1) non-stationary variables. However nearly half also may be considered I(2) variables also since their first differences turned out to be near

non-stationary. The second part of the two step Engle and Granger (1987) method consists of testing that some linear combination of the ADF tested variables is I(0). This was done for each country using an OLS regression of equation (43). The residuals from this regression were tested for non-stationarity. All rejected the null that $\hat{u}_t$ was I(1). These results are not shown.
To test for cointegration, in our panel, using panel methods we use the panel cointegration test outlined in Kao, Chiang and Chen (1999) (KCC). This test is for homogeous panel cointegration and is developped in Kao (1999) and highlighted in McKoskey and Kao (1999), Kao and Chiang (1998) and Kao, Chiang and Chen (1999). Phillips and Moon (1999) and Pedroni (1995), (1997) analyse similar models for cointegration tests in homogeneous panels. Dickey and Fuller (1981) (DF) type tests are calculated from the estimated residuals of the DOLS regression of equation (50) using

$$\hat{u}_{it} = \gamma \hat{u}_{it-1} + \pi_{it}. \tag{96}$$

where π_{it} is white noise. To test the null hypothesis of no cointegration the null is written as $H_0 : \gamma = 1$. The OLS estimate, $\hat{\gamma}$, of γ can be given as

$$\hat{\gamma} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{u}_{it} \hat{u}_{it-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{u}_{it}^2}. \tag{97}$$

KCC construct four DF type tests using $\hat{\gamma}$, we shall use only the first two given by

(a) $DF_\gamma = \frac{T\sqrt{N}\hat{\gamma} + 3\sqrt{N}}{\sqrt{10.2}}$,

(b) $DF_t = \sqrt{1.25} t_\gamma + \sqrt{1.875} N$.

Here t_γ is the t-statistic of $\hat{\gamma}$. The asymptotic distribution of DF_γ and DF_t converge to a standard normal distribution $N(0, 1)$. Our computed DF_γ and DF_t statistics are 4.0207 and 192.604, respectively. Both are statistically significant at the 5% and 1% levels with critical values of 1.96 and 2.58. Hence we unanimously reject the null hypothesis of no cointegration.
In Table 1 the preliminary analysis shows a quite good estimated fixed effects specification but as indicated by our Chow test we also need to consider the heterogeneous panel case. Starting with the ordinary t-statistics of column three, the intercepts for each country, modelled by the dummy variable method, are very strongly significant at the 5% level. All the t-statistics are over 10 in value and in a few cases t-statistics of over 150 are found. As for the expected signs and significance of the PPP regression parameters, they are very close to the joint proportionality and symmetry assumption of

$\beta_1 = -\beta_2 = 1$ and very strongly significant with t-values of around 100 each. However as shown by our LM_1 test there is serial correlation in the errors of an unspecified form, perhaps AR(1), hence these t-statistics are biased and not very useful for inference. In columns four and five we have the results obtained using the HAC standard error estimators described above. These correctly account for the serial correlation structure in the errors by using estimates of the long run variance of u_{it} in their computations. On average the HAC1 and HAC2 standard errors are about three times the size of the ordinary standard errors. These Newey and West estimators are the usual estimators refered to when estimating cointegrating regressions because of the endogeniety and serial correlation problems already described. The estimates using the HAC estimates are quite good with all the intercepts and levels regressors strongly significant at the 5% level, only now a few of the differenced regressors are insignificant. On the whole, the results seem to indicate that the the performance of the parametric DOLS estimator, in large samples, is quite close to the non-parametric FMOLS estimator. Indeed it may be prefered on account of its ease of use, in that there are no complicated prior adjustments needed for the OLS estimator. In order to apply the Cochrane-Orcutt serial correlation corrections, all the leads and lags were kept when estimating the pooled DOLS model. In Table 2 we have the DOLS estimation results for the group-mean panel cointegration model. These are much more conservative than the pooled panel estimates. Both the regressor point estimates, in columns two and four, and the t-statistics, in columns three and five, are consistently lower than those of their pooled panel counterparts. Sixteen of nineteen countries had domestic price estimates below unity and fourteen of nineteen for foreign prices. This indicates some underprediction in the model. The high R^2 values, in column seven, indicate that we still have good fits of the data. All leads and lags of the differenced regressors were chosen so that they were significant at least at the 10% significance level. In column six we have the leads and lag structures chosen for each country. Overall we can conclude that the pooled DOLS panel regression gives good support to the strong PPP hypothesis. Whilst the group-mean DOLS panel regression gives good support to the weak PPP hypothesis. This is not surprising considering that the former consists of a homogeneous panel and the later a heterogeneous one. In fact due to the rejection of the poolability assumption by our Chow test, the performance of the DOLS group-mean estimator becomes equally important as that of the pooled DOLS estimator.

In Tables 3 and 4 we have the results of the simple bootstrap Monte Carlo simulations. The simple bootstrap uses the regression of Table 1 columns two and three in the simulations. Even though the assumption of i.i.d. errors, necessary for the simple bootstrap, was weakened by assuming simple stationarity of the disturbance term, the bootstrap performs remarkably well here due probably to a number of facts. (i) The sample size is very large, thus even assuming weak stationarity, we have enough observations for any asymptotic independence assumption of the error term to be effective. (ii) We use an efficient DOLS estimator which has all the desirable properties of asymptotic normality and unbiasedness and also for its test statistics. In Table 3 we have the bootstrap pooled DOLS confidence intervals at the 80% nominal level and at the 90% nominal level in Table 4. All the bootstrap methods perform very well and are strong competitors for the asymptotic method. It should be noted that, throughout the study, the BC and BC_a methods gave the same results as the accelerator constant computed was close to zero. In terms of smallest coverage error overall the percentile method performs best here. At both the nominal levels considered the percentile method has the smallest coverage error for the coefficients of both foreign and domestic prices. In terms of the shortest interval, here the asymptotic method lead to the smallest overall intervals. The asymptotic method had the smallest intervals for all coefficients except for domestic prices at the 90% nominal level.

In Tables 5 and 6 we have the results of the pairs bootstrap Monte Carlo simulations for the group-mean DOLS estimator. The pairs bootstrap for the group-mean DOLS estimator uses the regressions of Table 2 in the simulations. With the group-mean DOLS estimates we have the percentile method again providing the smallest errors in coverage probabilities in all coefficients at both the 80% and 90% nominal levels. The BC and BC_a provide the shortest confidence intervals for domestic prices at the 80% nominal level. Whilst the bootstrap-t method has the shortest intervals for foreign prices at the 80% nominal level and for both domestic and foreign prices at the 90% nominal level. What is interesting here is the poor performance of the asymptotic method, with very large, confidence intervals (nearly twice those of the percentile methods) and errors in coverage probabilities.

Finally in Tables 7 and 8 we have the results of the pairs bootstrap Monte Carlo simulations for the pooled DOLS estimator. The pairs bootstrap for the pooled DOLS estimator uses the regression of Table 1, columns two and four, in the simulations. Here again the percentile method performs best with the smallest coverage errors for all coefficients at both nominal levels. Here also the asymptotic method provides the shortest intervals for all coefficients at both nominal levels. It has intervals approximately half the size of those of the bootstrap-t method. However it does so at the cost of very large coverage errors. Whereas the bootstrap-t method performs quite reasonably at both nominal levels. Overall the effect of correctly accounting for serial correlation in the residuals of the cointegration regression and using the pairs bootstrap rather than the simple bootstrap method, is that both the group-mean DOLS and pooled DOLS bootstrap estimators have much larger confidence intervals, due to the larger standard errors. However coverage errors are roughly the same. The asymptotic method is also shown to be quite unstable in this panel data cointegration regression. First with the group-mean DOLS estimator it overcovers by as much as $10\% - 20\%$ at both nominal levels. Then with the pooled DOLS estimator it undercovers by as much as $10\% - 20\%$ at both nominal levels. Whilst all the bootstrap methods deliver consistently small coverage errors of less than 5%. The pairs bootstrap method hence is shown to be remarkably accurate and efficient both with the group-mean and pooled DOLS estimators.[10] Overall the percentile methods (including BC and BC_a) seem to be best in delivering optimal confidence intervals, in terms of shortest intervals and smallest coverage errors. Our results do coincide with some of the findings in the bootstrap literature. However few (or no) applications exist, in the econometric literature, of bootstrap confidence interval studies with cointegrated (panel) regressions. Kilian (1999) conducts a Monte Carlo analysis of the coverage accuracy and average length of bootstrap confidence intervals, for impulse response estimators. He finds the central percentile and percentile-t intervals both shorter and more accurate than equal-tailed and symmetric percentile-t intervals. Kilian (1999) makes very similar assumptions, as made in this paper, in that he applies the simple bootstrap to the residuals of an AR(1) regression where the disturbance $u_t \sim$ NID(0,1). That is u_t is only normally identically distributed with some time dependency allowed between errors. Kilian conducts a sensitivity analysis, with respect to sample size, to see the effects of varying the sample sizes from $T = 100, 200, 500$ to $2,000$. He finds that some methods (eg percentile-t) perform poorly in smaller samples, but the performance of all methods improve with increasing sample sizes. For the very large sample size (2,000) all bootstrap methods have excellent coverage. Another econometric application of the bootstrap has come from Kazimi and Brownstone (1999) who provide bootstrap confidence bands for shrinkage OLS estimators. They find that of the numerous methods they consider that the BC_a methods perform best and generate the tightest bands with coverage close enough to the nominal level. Next the percentile method used with the single bootstrap is best. Whilst the symmetric percentile with the double bootstrap also performs well. The standard or asymptotic method

[10]The results shown use the HAC1 standard error estimates in computations. Those using the HAC2 estimates were much more inferior with all methods except the percentile methods

does not perform particularly well. In the general statistical literature the bootstrap studies tend to be more theoretical and rigorous. Diciccio and Romano (1988) provide examples of the percentile, BC, BC_a and percentile-t methods for the confidence intervals of a correlation coefficient and an exponential mean. They find that the BC and BC_a improve on the percentile method in terms of accuracy. Finally Efron (1987) introduces his corrections to the percentile method for the central $(1-\alpha)$ BC and BC_a intervals in a theoretical study. He shows the BC_a interval is nearly identical to the exact interval. Whilst the BC interval only partially improves on the standard or asymptotic interval. This last interval has the largest length, followed by the BC, and then the BC_a and exact methods.
Returning to our results we have illustrated the very high performance of the bootstrap methods when used in panel cointegration regressions. The pairs bootstrap outperforms the simple bootstrap since all information in the cointegration model is taken into account when constructing bootstrap samples. This lead it to having the most accurate coverage probabilities and close to the shortest confidence intervals. Also the bootstrap methods that excel are the percentile methods that use only the bootstrap statistic's distribution in their computations. However the asymptotically pivotal bootstrap-t method also performs well. We have shown also that any serial correlation detected in the residuals of the cointegration regression should be corrected with an application of the MBB, or recursive bootstrap methods. This has the effect of improving the performance of all estimators and bringing the asymptotic and bootstrap-t methods (which rely on accurate estimates of the regression standard errors) in line with the percentile methods.

8 Conclusion

In this study we have shown that the bootstrap can effectively be used to construct confidence intervals in panel data cointegrated regressions. When efficient DOLS estimators such as the pooled (Kao and Chiang (2000)) DOLS or group-mean (Pedroni (2001)) DOLS estimators are used, in a large macro panel cointegration regression, one can use the pairs bootstrap method, with modifications where necessary, to provide bootstrap estimators that are as least as efficient as estimators based on first-order asymptotic theory. Using our optimal confidence interval criterion we found the bootstrap percentile method provided both the shortest confidence intervals and the smallest coverage errors.