PPP over a century: Cointegration and structural change

Ekaterini Panopoulou

National University of Ireland, Maynooth[*]

February 2006

Abstract

The purpose of this paper is to investigate the ability of parameter instability tests in regressions with I(1) processes to discriminate between changes in the cointegrating relationship and changes in the marginal distribution of the regressors. Using annual data for the G-7 countries and the Purchasing Power Parity, we conclude that the regression coefficient between the price level differential and the exchange rate has indeed remained stable during the 20th century and find ample evidence supporting the PPP.

[*]*Correspondence to*: Ekaterini Panopoulou Department of Economics, National University of Ireland, Maynooth, Co. Kildare, Republic of Ireland. E-mail: apano@nuim.ie. Tel: 00353 1 7083793. Fax: 00353 1 7083934.

1. Introduction

Testing for regression parameter constancy has a long history in economics. The main body of the literature, however, is devoted to regressions involving stationary series (see Perron, 2005, for a literature review and the references therein). Since Engle and Granger's seminal paper in 1987, the issue of estimating cointegration parameters has received considerable attention in both the theoretical and applied econometrics literature. Consequently, parameter constancy tests valid for cointegrating relationships were developed. Hansen (1992) was the first to extend structural break tests to regressions involving integrated variables. These tests are derived as Langrange- Multiplier (LM) tests in correctly specified likelihood problems. The author, using the fully modified estimator of Phillips and Hansen (1990), derives the limiting distribution of the respective test statistics under the null of coefficient constancy, which follow asymptotically non-standard, but nuisance free distributions.

The purpose of this paper is to investigate the ability of these parameter instability tests in regressions with I(1) processes to discriminate between changes in the cointegrating relationship and changes in the marginal distribution of the regressors.[1] The evaluation takes place within a bivariate cointegrating framework of the most popular parity relationship in International Economics, the Purchasing Power Parity (PPP). Using annual data for the G-7 countries, we find that the regression coefficient between the price level differential and the exchange rate has indeed remained stable during the 20th century. Our results suggest that when a single break occurs in the process that drives the regressor, while the parameter of interest remains stable, the tests display severe size distortions, leading to an overrejection of the null. By means of Monte Carlo simulations, we obtain small sample critical values for the tests along with the empirical distribution of some commonly used cointegration estimators. On the whole, we establish the stability of the cointegrating coefficient and find ample evidence supporting PPP in the G-7 countries over the last century.

The rest of the paper is organised as follows: Section 2 briefly reviews the stability tests. Section 3 reports the empirical and Monte Carlo results and section 4 concludes the paper.

2. Structural stability tests

The model employed by Hansen (1992) is a standard multiple regression cointegrating model: $y_t = Ax_t + u_{1t}$, $t = 1, 2, ...n$, where $x_t = (x'_{1t}, x'_{2t})'$ given by the equations: $x'_{1t} = k_{1t}, x'_{2t} = \prod_1 k_{1t} + \prod_2 k_{2t} + x^0_{2t}$ and $x^0_{2t} = x^0_{2t-1} + u_{2t}$. The elements of u_t and k_t are a sequence of mean zero random vectors and nonegative integer powers of time, respectively. Associated with the previous system are the long-run covariance matrix, Ω, and the one-sided covariance matrix, Λ, defined as $\Omega = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{j=1}^{n} E(u_j u'_t)$ and $\Lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{j=1}^{t} E(u_j u'_t)$. We partition Ω and Λ comformably with u and have $\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$ and $\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}$. Then $\Omega_{1.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$ represents the long-run variance of u_{1t} conditional on u_{2t} and $\Lambda_{21}^+ = \Lambda_{21} - \Lambda_{22}\Omega_{22}^{-1}\Omega_{21}$ the bias due to the endogeneity of the regressors.

---

[1]Hansen (2000) shows that in a stationary environment the asymptotic distribution of these tests is affected by changes in the marginal distribution of the regressors.

Hansen (1992) proposed three tests for parameter constancy in the context of cointegration, namely the $SupF$, $MeanF$, and L_c. They all require an asymptotic mixed normal estimate of A and have the same null hypothesis but differ in their choice of alternative hypothesis. The null hypothesis of the three tests is that the coefficient A is constant and the timing of the break is unknown. The test statistic employed by the first two tests is

$$F_{nt} = vec(S_{nt})'(\widehat{\Omega}_{1\cdot2} \otimes V_{nt})^{-1} vec(S_{nt}) = tr\{S_{nt}' V_{nt}^{-1} S_{nt} \widehat{\Omega}_{1\cdot2}^{-1}\}$$

where $S_{nt} = \sum_{i=1}^{t} \hat{s}_i$, $\hat{s}_t = \left(x_t \hat{u}_{1t}^+ - \binom{0}{\hat{\Lambda}_{21}^+}\right)$, $V_{nt} = M_{nt} - M_{nt} M_{nn}^{-1} M_{nt}$, and $M_{nt} = \sum_{i=1}^{t} x_i x_i'$. The main difference is the treatment of alternative hypotheses.

The first test models A as undergoing a single structural break at unknown time t and the alternative hypothesis is $H_1 : A_1 \neq A_2$. This test statistic is simply $supF = \sup_{[t/n]\in T} F_{nt}$, $[t/n] \in T$, where T is some compact subset of $(0,1)$ and $[\cdot]$ denotes integer part. The second and third test model the parameter A_t as a martingale process: $A_t = A_{t-1} + \varepsilon_t$; $E(\varepsilon_t/T_{t-1}) = 0$ and $E(\varepsilon_t \varepsilon_t') = \delta^2 G_t$, where $G_t = (\widehat{\Omega}_{1\cdot2} \otimes V_{nt})^{-1}$. Then the null hypothesis can be formulated as a constraint that the variance of the martingale differences is 0: $H_0 : \delta^2 = 0$. The alternative hypothesis is $H_2 : \delta^2 > 0$, with test statistic $MeanF = \frac{1}{n^*} \sum_{[t/n]\in T} F_{nt}$, where $n^* = \sum_{[t/n]\in T} 1$ and $[t/n] \in T$. The third alternative is $H_3 : \delta^2 > 0$, with test statistic $L_c = tr\{M_{nn}^{-1} \sum_{t=1}^{n} S_t \widehat{\Omega}_{1\cdot2} S_t'\}$.

The $supF$ dates back to Quandt (1960) and entails the choice of the region T. Following Andrews (1993) and Hansen (1992), we select $T = [.15, .85]$. Some trimming is also required for the $meanF$, which is simply the average F_{nt} test. The easier of all to compute is the L_c test, which requires no trimming. The asymptotic distributions of the test statistics are non standard and depend on the stochastic process of the regressors. In this respect, asymptotic critical values were calculated by Monte Carlo simulations (Tables 1-3, Hansen, 1992). Moreover, the tests are dependent on the estimation procedure of the cointegrating relationship. In this study, we calculate the tests via the Fully Modified (FM-S) estimation technique of Phillips and Hansen (1990), which in turn requires the selection of a kernel and the determination of the bandwidth. We employ the Quadratic Spectral kernel and select the bandwidth parameter by applying the Andrews (1991) data-dependent procedure. Moreover, the "prewhitened" version of FM-S (FM-PW) which filters the error vector $\widehat{\mathbf{u}}_t$ prior to estimating Ω and Δ is employed.

3. Empirical results and Monte Carlo simulations

3.1. Data

Our data set consists of annual data on the bilateral exchange rate over USD and the WPI of the G-7 countries (excluding Germany).[2] The sample covers a century of data, namely the period 1800-1999.[3] In particular, we focus on the PPP relationship, linking nominal exchange rates to price differentials and requiring one-to-one adjustment of the former to the latter. Formally, PPP

---

[2] Data for Germany are not reliable due to the unification process.

[3] The data are obtained from Global Financial Data, Annual Worksheets. The countries considered are the following (with the first year of the sample in brackets): Canada (1900), France (1900), Italy (1910), Japan (1900), UK (1900), US (1900).

can be expressed by the following regression $s_t = a + b(p_t^* - p_t) + v_t$, where s_t and $p_t^* - p_t$ is the nominal exchange rate and the price gap between two markets.[4] The null hypothesis to be tested can take the form: PPP holds $\Leftrightarrow b = 1$. Usually, the literature agrees on the fact that s_t and $(p_t^* - p_t)$ are cointegrated processes. On the other hand, estimates of b appear to be significantly different from unity, thus casting doubts on the parity. As a first step, we carried out typical unit root tests and cointegration tests. On the whole, the results (not reported) broadly confirm those of other studies, i.e. nominal exchange rates and price differentials seem to be I(1) and cointegrated processes.

3.2. Structural stability tests

After establishing that the involved variables are integrated and cointegrated, we proceed with testing for the stability of the cointegration parameter since rejecting the null of stability of the cointegrating coefficient constitutes an empirical failure of the PPP and probably estimation has to be done in different subsamples.

Table 1 reports the results of the stability tests. On the basis of the *supF* test, we reject the null of stability in all countries with the break estimated at around the end of World war II. The other two tests provide mixed evidence on the issue with results being more robust for Canada, Italy and Japan.[5] Employing the standard or the prewhitened version of the FMLS estimator does not in general lead to conflicting results, with the exception of the timing of the break in Italy and UK, which is shifted by 5 and 8 periods, respectively.

[INSERT TABLE 1]

3.3. Monte Carlo simulations

Here, we examine whether the presence of a break in the marginal distribution of the regressors affects the size of the tests which probably erroneously detect it as a coefficient break. To do this, we assume that s_t and $(p_t^* - p_t)$ are generated by the following bivariate triangular DGP:

$$s_t = a + b(p_t^* - p_t) + u_{1t} \tag{1}$$

$$\Delta(p_t^* - p_t) = u_{2t} \tag{2}$$

We further assume that $\mathbf{u}_t = [u_{1t}, u_{2t}]'$ is an I(0) process and follows a VAR(1) process, i.e $\mathbf{u}_t = A_t \mathbf{u}_{t-1} + \mathbf{e}_t$:and $\mathbf{e}_t \sim NIID(0, \Sigma_t)$. The model parameters are calibrated from the exchange rate and relative price data from the countries at hand. In particular, the parameter of interest, b, is set equal to one for the whole sample, while the covariance matrix Σ_t and the transition matrix A_t is allowed to undergo a single break at the time indicated by the *sup F* test (see Table 1, FM-PW).[6] We consider a sample size of 150 observations and replicate simulations 5000 times. To account for the effect of the initial conditions, the first 50 observations are discarded. The percent rejections of the null for a nominal 5% size along with the average estimated timing of the

---

[4] All variables in logs.

[5] Following Hansen (1992) the cut-off point p-value is set at 0.20.

[6] The parameter values of the covariance matrix Σ_t and the transition matrix A_t for the two subsamples and the five countries under scrutiny are not reported for brevity but are available upon request from the author.

break and its attendant standard error for the three tests and the countries at hand are reported in Table 2. We also calculate new critical values corresponding to an empirical size of 5% and employ them to assess whether stability is rejected or not. These results along with the desicion of the tests are reported in Table 3.

[INSERT TABLE 2 & 3]

The information content in these Tables may be summarised as follows:

(i) All the tests are oversized for all countries with the exception of the Lc test (FM-PW version) for Canada. Size distortions in some cases reach 100%. For example, the size of both the $meanF$ and $sup\ F$ for Italy is 99%.

(ii) The Lc test turns out to be the least distorted. With the exception of Italy, the respective values for the remaining countries range from 3.6% to 13%.

(iii) Employing the prewhitened version of FMLS does not necessarily lead to size gains. On the contrary, the size of the tests increases considerably for France, Italy and Japan.

(iv) The average timing of the break is estimated with a large standard deviation, which can reach 23 years in some cases.

(v) On the basis of the small sample critical values of the tests (Table 3), we overall cannot reject the null of stability of the cointegration coefficient. The only cases we reject the null of stability is on the grounds of the Lc test for Canada and the SupF test for France.

Overall our results suggest that during the last century the marginal distributions of exchange rates and price differentials have witnessed a structural break, while the parameter of interest has indeed remained constant. Next, we estimate the cointegrating parameter and test whether it is equal to one for the countries at hand.

3.4. Cointegration estimators

The cointegration estimators we employ are the Ordinary Least Squares (OLS), the Autoregressive Distributed Lag (ADL) (Pesaran and Shin, 1999), the Johansen's (JOH) maximum likelihood (1988, 1991) and the Fully Modified least squares estimators (FM-PW) (Phillips and Hansen, 1990) estimators. The estimates of b are reported in Table 4 along with the associated standard errors and the t-tests for the null hypothesis of interest $b = 1$ for all the countries under consideration. We also test the hypothesis of interest based on the 2.5% ($t_{0.025}$) and the 97.5% ($t_{0.975}$) points in the empirical distributions of the relevant t-statistics through Monte Carlo simulations.[7] These simulated critical values along with the empirical sizes of the t-tests for nominal sizes of 5% are reported in Table 5.

[INSERT TABLE 4&5]

On the basis of asymptotic critical values, PPP survives the empirical evidence only for Canada and Italy irrespective of the estimator employed. For the remaining countries evidence is mixed as we cannot reject the null of a unit coefficient for France based on JOH and for the UK based on the ADL and FM-PW estimators. Naturally, the performance of the estimators is affected by the presence of a break. Our simulations suggest that the distributions of the relevant t-statistics become less leptokurtic and in most cases shift to the left (Table 5). This leads to significant size increases, which for the asymptotically efficient estimators range from 9.6% to

---

[7]The design of the Monte carlo experiment is similar to the one in Section 3.3.

44.4% for a nominal 5% size. Consequently, employing critical values leads to fewer rejections of a unit coefficient hypothesis. Specifically, we find evidence in favor of PPP for all the countries at hand when small sample critical values and the asymptotically efficient estimators are employed.

4. Conclusions

In this paper we have examined the effects of a break in the marginal distribution of integrated regressors on the size of some commonly used stability tests for the cointegration coefficient, namely the $SupF$, $MeanF$, and L_c (Hansen, 1992). By means of Monte Carlo simulations we showed that the tests erroneously detect a break in the cointegrating coefficient when the break actually occurs in the marginal distribution of the regressors. These issues were addressed in an empirical framework, specifically the Purchasing Power Parity for the G-7 countries for which we found ample evidence.

References

[1] Andrews, D.W.K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 817-858.

[2] Andrews, D.W.K., 1993, Tests for parameter instability and structural change with unknown change point, Econometrica 61, 821-856.

[3] Engle, R.F. and C.W.J. Granger. 1987, Cointegration and error correction representation, estimation and testing, Econometrica 55, 251-276.

[4] Hansen, B.E., 1992, Tests for parameter instability in regressions with I(1) processes, Journal of Business and Economic Statistics 20, 45-59.

[5] Hansen, B.E., 2000, Testing for structural change in conditional models, Journal of Econometrics 97, 93-115.

[6] Johansen, S., 1988, Statistical analysis of cointegrating vectors, Journal of Economic Dynamics and Control, 12, 231-254.

[7] Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica, 59, 1551-1580.

[8] Perron, P., 2005, Dealing with structural breaks, Palgrave Handbook of Econometrics, Vol.1: Econometrics Theory.

[9] Pesaran H.M. and Y. Shin (1999), An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis, Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, chapter 11. Cambridge University Press, Cambridge.

[10] Phillips, P.C B and Hansen, B.E, 1990, Statistical inference in instrumental variables regression with I(1) processes, Review of Economic Studies 57, 99-125.

[11] Quandt, R. , 1960, Tests of the hypothesis that a linear regression system obeys two separate regimes, Journal of the American statistical Association 55, 324-330.

Table 1: Structural Stability Tests

Country	Canada		France		Italy		Japan		UK	
	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW
L_c	1.027	0.665	0.177	0.150	0.278	0.324	0.308	0.275	0.327	0.116
	(0.01)	(0.01)	(0.20)	(0.20)	(0.17)	(0.13)	(0.14)	(0.18)	(0.13)	(0.20)
$MeanF$	8.204	5.302	2.647	2.199	4.810	11.922	9.677	11.698	2.904	2.732
	(0.01)	(0.03)	(0.20)	(0.20)	(0.04)	(0.01)	(0.01)	(0.01)	(0.20)	(0.20)
$SupF$	19.429	14.918	19.753	18.130	33.754	88.843	8.892	34.42	4.376	7.448
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$Break$	1946	1946	1945	1945	1934	1939	1948	1948	1954	1946

Note: p-values in parentheses below tests.

Table 2: Size of Structural Stability Tests

Country	Canada		France		Italy		Japan		UK	
	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW
L_c	9.10	3.60	5.85	6.45	55.00	81.75	6.95	12.50	12.25	5.35
$MeanF$	16.40	7.50	8.70	9.50	99.15	100	29.15	42.45	39.55	20.80
$SupF$	18.00	9.10	6.70	8.55	98.80	100	21.35	36.55	37.95	18.10
$Break$	35	36	27	26	46	47	45	45	45	39
$(s.e.)$	(23)	(23)	(20)	(20)	(14)	(14)	(19)	(19)	(21)	(23)

Note: Empirical size for a nominal 5% size.

Table 3: Monte Carlo Critical Values

Country	Canada		France		Italy		Japan		UK	
	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW	FM-S	FM-PW
L_c	0.729	0.507	0.611	0.648	1.813	2.845	0.622	0.765	0.745	0.55
	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$MeanF$	8.619	5.328	5.431	5.656	83.35	150.88	10.29	16.79	7.424	4.888
	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$SupF$	28.47	16.27	13.23	14.376	277.75	492.53	26.49	44.21	19.04	10.39
	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Note: "Yes" indicates non-rejection of the null of coefficient constancy and "No" indicates rejection of the null.

Table 4: Estimation results

Country	Canada		France		Italy		Japan		UK	
Estimator	b	t_{stat}	b	t_{stat}	b	t_{stat}	b	t_{stat}	b	t_{stat}
OLS	0.914	-0.694	1.040	2.938	0.963	-1.663	0.942	-4.425	0.783	-4.227
	(0.124)		(0.014)		(0.023)		(0.013)		(0.051)	
ADL	1.056	0.241	1.043	2.243	0.978	-0.731	0.927	-2.114	0.839	-1.508
	(0.234)		(0.019)		(0.030)		(0.034)		(0.107)	
JOH	1.174	0.994	1.122	0.705	0.995	-0.208	0.952	-2.400	0.754	-2.625
	(0.175)		(0.173)		(0.024)		(0.020)		(0.094)	
$FM-PW$	1.064	0.405	1.044	3.143	0.997	-0.143	1.049	2.631	1.295	1.815
	(0.158)		(0.014)		(0.021)		(0.019)		(0.163)	

Note: Standard errors in parentheses below coefficients.

Table 5: Small-sample performance of cointegration estimators

	Canada		France		Italy		Japan		UK	
Estimator	$t_{0.025}$	$t_{0.975}$	$t_{0.025}$	$t_{0.975}$	$t_{0.025}$	$t_{0.975}$	$t_{0.025}$	$t_{0.975}$	$t_{0.025}$	$t_{0.975}$
OLS	-13.221	7.231	-2.780	2.266	-6.672	8.538	-4.051	3.839	-4.781	4.908
	(67.80)		(14.05)		(59.20)		(30.30)		(40.30)	
ADL	-4.448	3.044	-2.633	2.255	-3.636	2.950	-2.491	2.549	-2.158	2.812
	(24.10)		(10.60)		(20.05)		(9.90)		(11.40)	
JOH	-5.221	4.143	-2.499	2.378	-2.820	2.055	-3.250	3.445	-2.816	2.967
	(44.35)		(10.15)		(9.60)		(22.10)		(16.15)	
$FM-PW$	-2.010	2.665	-2.576	2.623	-4.433	1.874	-3.024	3.211	-2.393	2.756
	(10.85)		(13.90)		(25.10)		(17.90)		(12.50)	

Note: Empirical size of the t-tests for a 5% empirical size in parentheses.