Cointegration Tests of Purchasing Power Parity

Frederick H. Wallace
Universidad de Quintana Roo
Chetumal, Quintana Roo
México

Revised
October 27, 2007

I thank Alan Taylor for providing the data used in this study and Walt Enders for providing computer code and recent versions of his working papers cited herein. Gary Shelley made a number of useful suggestions. Financial support from the National Council for Science and Technology (CONACYT) of Mexico is gratefully acknowledged.

This is a very preliminary draft describing ongoing research. Please do not quote.

Abstract: Cointegration Tests of Purchasing Power Parity: Preliminary Evidence

JEL Codes: C22, F31
Keywords: Cointegration, purchasing power parity

In recent work Im, Lee, and Enders (2006) use stationary instrumental variables to test for cointegrating relationships. The advantage of their test is that the t-statistics are asymptotically standard normal and the familiar critical values of the normal distribution may be used to assess significance. Thus the test avoids various complications arising in regressions with integrated variables. Using the data set developed by Taylor (2002) to test for purchasing power parity, the ILE test is compared to three single equation alternatives: An error correction model, autoregressive distributed lag model, and the Engle-Granger two step procedure. The empirical evidence from the Taylor data raises a number of questions concerning the new test and the single equation alternatives.

Introduction

The hypothesis of purchasing power parity (PPP) has been the focus of much empirical work. Simply stated, PPP says that the price of a market basket of (traded) goods is the same everywhere in terms of a common currency. The concept is important because theories in open economy macroeconomics typically imply PPP as a long run equilibrium condition. A partial list of techniques used in such empirical work includes single equation unit root tests, variance ratio tests, cointegration studies, and panel unit root tests. Some of these methodologies have been adapted for use as nonlinear procedures. Underlying the PPP hypothesis is the law of one price (LOOP) which indicates that the price of a (traded) good is the same in all locations in terms of a common currency. Rather than focus directly on PPP, numerous studies have examined the LOOP with the idea that support for the law of one price implies support for PPP. Sarno and Taylor (2002) provide a thorough review of the PPP and LOOP literature.

The purpose of this paper is to compare the results from standard, single equation cointegration tests of purchasing power parity with those from new tests developed by Im, Lee, and Enders (2006), henceforth ILE. Tests are carried out using the data set containing 100+ annual observations for twenty countries constructed by Taylor (2002).[1] Applying the Elliot, Rothenberg, and Stock (ERS, 1996) unit root test to transformed (demeaned or detrended) data, Taylor finds support for PPP with respect to the United States in eighteen of nineteen series. Only data for Japan fail to indicate PPP for either transformed series. When purchasing power parity is tested with respect to a world market basket, Taylor finds evidence in favor of the hypothesis using demeaned or detrended data in nineteen of the twenty series. Data for Canada fail to reveal any support for PPP. Lopez, Murray, and Papell (2005) argue that Taylor's results can be

[1] Argentina, Australia, Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the United States.

attributed to the selection of suboptimal lag length in his unit root tests. Employing optimal lag length selection criteria, they conclude that the data support PPP with respect to the US in just nine of sixteen countries.[2] Instead of relying on unit root tests Wallace and Shelley (2006) apply the Fisher-Seater test with bootstrapped errors to the Taylor data and conclude that PPP holds for at least twelve of nineteen countries with respect to the United States.

Methodology

Standard tests for cointegration have asymptotic distributions which are not standard normal and which may depend a nuisance parameter. Pesavento (2007), for example, evaluates the analytical local power of five residual-based cointegration tests and shows that test power is dependent on the nuisance parameter R^2. She finds that power is low in all five tests when the nuisance parameter is large. Im, Lee, and Enders describe single equation cointegration tests in which stationary variables are used as instruments. Hence, there are no nuisance parameters and the asymptotic distributions are standard normal. A brief description of the ILE methodology, using their notation, is provided to assist in understanding the empirical results. For more detailed treatment see their working paper.

Starting with a VAR(p) model in which the variables are cointegrated, ILE derive an error correction model (ECM) of the form given by equation (1)

$$\Delta y_{1t} = (d_{11} + d_{12}t) + \delta_1 z_{t-1} + \phi \Delta y_{2t} + C_{11}\Delta y_{1t-1} + C_{12}\Delta y_{2t-1} + \nu_t \qquad (1)$$

where y_{it}, t = 1,2, …T, i = 1,2, are I(1) processes, the d_t are deterministic terms, $z_{t-1} = y_{1t-1} - \beta y_{2t-1}$, and ν_t is a linear combination of the normally distributed and independent errors of the original VAR.[3] The vector error correction model derived

---

[2] They eliminate Argentina, Brazil, and Mexico from their study.

[3] ILE assume normality for convenience and point out that the assumption does not affect the asymptotic results.

from the original VAR reduces to a single equation if y_{2t} is weakly exogenous as will be assumed for this study. The null (of no cointegration) and alternative hypotheses are given by $H_0: \delta_1 = 0 \qquad H_1: \delta_1 < 0.$

Alternatively, the ECM can be rewritten as the autoregressive distributed lag (ADL) in equation (2)

$$\Delta y_{1t} = (d_{11} + d_{12}t) + \delta_1 y_{1t-1} + \gamma y_{2t-1} + \phi \Delta y_{2t} + C_{11}\Delta y_{1t-1} + C_{12}\Delta y_{2t-1} + \nu_t \qquad (2)$$

with the same null and alternative as the ECM test.

The Engle-Granger (EG) test, of course, is based on the two step procedure whereby i) y_{1t} is regressed on y_{2t} using ordinary least squares and ii) the estimated residuals are tested for a unit root as in equation (3),

$$\Delta\left(y_{1t} - \hat{\beta} y_{2t}\right) = \delta_1 \left(y_{1t-1} - \hat{\beta} y_{2t-1}\right) + C(L)\Delta\left(y_{1t} - \hat{\beta} y_{2t}\right) + u_t \qquad (3)$$

where $\hat{\beta}$ is the estimated vector of parameters from $y_{1t} = (d_{11} + d_{12}t) + \hat{\beta} y_{2t} + \varsigma_t$, with d_{11} as an (optional) constant, t as an (optional) time trend, and the $C(L)\Delta\left(y_{1t} - \hat{\beta} y_{2t}\right)$ are lags of the estimated residuals. The null and alternative hypotheses are

$$H_0 : \delta_1 = 0 \qquad H_1 : \delta_1 \neq 0$$

Weak exogeneity is not necessary for the EG test. ILE note that the EG test is not plagued by the nuisance parameter problem but that it can lose power under some conditions. In all three tests, δ_1 has a nonstandard distribution under the null.

ILE suggest using instrumental variables (IV) to address the problems occurring in the three single equation cointegration tests described above. Specifically, they suggest defining the instrumental variable w_t given by:

- $w_t = z_{t-1} - z_{t-m-1}$ for z_{t-1} in (1)
- $w_t = \left[\left(y_{1t-1} - y_{1t-m-1}\right), \left(y_{2t-1} - y_{2t-m-1}\right)\right]$ for $\left(y_{1t-1}, y_{2t-1}\right)$ in (2)
- $w_t = \left(y_{1t-1} - \hat{\beta} y_{2t-1}\right) - \left(y_{1t-m-1} - \hat{\beta} y_{2t-m-1}\right)$ for $y_{1t-1} - \hat{\beta} y_{2t-1}$ in (3)

and m < T. ILE suggest increasing m when the errors are autocorrelated. A constant with or without trend may be added to each equation. ILE show that the t statistic for $\delta_1 = 0$ (t_{ECM}, t_{ADL}, or t_{EG}) in the equation with instruments has a standard normal distribution for a variety of specifications provided any other nonstationary variables are instrumented. Furthermore, they note that the estimated coefficient $\hat{\delta}_{1i}$ $\left(\hat{\delta}_{1ECM}, \hat{\delta}_{1ADL}, or\, \hat{\delta}_{1EG}\right)$ is consistent but does not follow a normal distribution.[4]

An unresolved issue in their tests concerns the optimal selection of m for which theory offers no resolution. They explore the use of different values of m in simulations. In a related paper Enders, Lee, and Strazicich (2007) suggest selecting the value of me that minimizes the sum of the squared residuals. Different values of m are used in the cointegration tests for purchasing power parity as a robustness check.

Data and Empirical Results

The Taylor data set consists of annual observations on nominal exchange rates and price indexes (usually a consumer price index) for the twenty countries listed in footnote 1. The nominal exchange rate is measured as units of foreign currency per US dollar. For each country the data span more than 100 years, ending in 1996. Based on augmented Dickey-Fuller, ERS, and the KPSS [Kwiatkowski et. al. (1992)] tests, Wallace and Shelley conclude that all nominal exchange rates and price indexes are integrated of order one, I(1).

If purchasing power parity holds relative to the price level in the United States, it can be expressed as equation (4)

$$f_t = p_t^F - e_t = \alpha + \beta p_t^{US} \tag{4}$$

[4] Again, see ILE for proofs and more detail.

where e_t is the log of the price of a US dollar in terms of the foreign currency, p_t^{US} is the log price level in the United States, p_t^F is the log foreign price level, while f_t is interpreted as the dollar denominated foreign price level. Absolute purchasing power parity implies the coefficient restrictions α = 0 and β = 1 but in practice, due to the use of price indices rather than actual measures of the price level, equation (4) with these restrictions rarely holds. But the basis of cointegration tests is that PPP implies the existence of a cointegrating relation between f_t and p_t^{US}. In terms of equations (1)-(3) the ECM, ADL, and EG cointegration tests for purchasing power parity can be written as equations (5)-(7), respectively.[5]

$$\Delta f_t = d_{11} + \delta_1 \left(f_{t-1} - \alpha - \beta p_{t-1}^{US} \right) + \phi \Delta p_t^{US} + \nu_t \tag{5}$$

The expression in parentheses in equation (5) is the error, lagged one period, from the estimation of equation (4), that is, the error correction term. The US price level, p_t^{US} is assumed to be weakly exogenous.

The ADL form of the model is

$$\Delta f_t = d'_{11} + \delta_1 f_{t-1} + \gamma' p_{t-1}^{US} + \phi \Delta p_t^{US} + \nu_t \tag{6}$$

where $d'_{11} = d_{11} - \delta_1 a$ and $\gamma' = \delta\beta$. For the ECM and ADL versions, the same null and alternative apply,

$$H_0: \delta_1 = 0 \qquad H_1: \delta_1 < 0.$$

The null implies the absence of a cointegrating relation between the US price level and the foreign dollar denominated price level. In other words, failure to reject the null would imply that PPP does not hold. Lagged values of f_t can be added to equations (5)

[5] Since the PPP relationship does not include a deterministic time trend; t is omitted from the empirical models.

and (6) in the case of serial correlation. Finally the Engle-Granger two step procedure involves testing for a unit root in the estimated residuals from equation (4).

$$\Delta\left(f_t - \hat{\alpha} - \hat{\beta}p_t^{US}\right) = \delta_1\left(f_{t-1} - \hat{\alpha} - \hat{\beta}p_{t-1}^{US}\right) + \sum_{i=1}^{j}\phi_i\Delta\left(f_{t-i} - \hat{\alpha} - \hat{\beta}p_{t-i}^{US}\right) + u_t \qquad (7)$$

Each of the single equation empirical models given by (5)-(7) is estimated in the form specified with results compared to its estimation using the instrumental variables w_t where

- $w_t = \left(f_{t-1} - \alpha - \beta' p_{t-1}^{US}\right) - \left(f_{t-m-1} - \alpha - \beta' p_{t-m-1}^{US}\right)$ for $f_{t-1} - \alpha - \beta' p_{t-1}^{US}$ in (5)
- $(w_{1t}, w_{2t}) = \left[\left(f_{t-1} - f_{t-m-1}\right), \left(p_{t-1}^{US} - p_{t-m-1}^{US}\right)\right]$ for $\left(f_{t-1}, p_{t-1}^{US}\right)$ in (6)
- $w_t = \left(f_{t-1} - \hat{\alpha} - \hat{\beta}' p_{t-1}^{US}\right) - \left(f_{t-m-1} - \hat{\alpha} - \hat{\beta}' p_{t-m-1}^{US}\right)$ for $\left(f_{t-1} - \hat{\alpha} - \hat{\beta}' p_{t-1}^{US}\right)$ in (7)

Estimation of the error correction model (ECM) given by equation (5) for each country yields t-statistics on the estimated error correction coefficient, $\hat{\delta}_1$, shown in column 2 of Table 1. Marginal significance levels (i.e. p values) were calculated using a program from Ericsson and MacKinnon (2002). The results suggest weak support for purchasing power parity, just five countries display estimated coefficients on the error correction term that are significant at the 5% level or better. However, it is likely that serial correlation affects the results, an issue addressed in the IV estimations reported later.

For comparison columns 3-6 of Table 1 show the t-statistics on $\hat{\delta}_1$ when the instrument, w_t t = 2,4,7,9 replaces the usual error correction term. Since ILE show that the distribution of the statistics shown in these four columns is standard normal, the critical values of -1.645 (5%) and -1.96 (1%) can be applied. Except in a few cases, the t-statistics on $\hat{\delta}_1$ are not significant, indicating a failure to support PPP, when using a small value (2 or 4) of m in constructing the instrument. Conclusions change somewhat with instruments using larger (7 or 9) values of m. For m = 7, the null of no

Table 1 t-Statistics from Estimation of Equation (5) Without and With Instruments

		Instrumental Variable			
Country	ECM	m = 2	m = 4	m = 7	m = 9
Argentina	-5.039**	-1.646	-4.384**	-1.787*	-2.877**
Australia	-2.026	0.986	-0.624	-1.164	-1.720*
Belgium	-3.160	0.546	-1.093	-1.288	-1.292
Brazil	-2.402	-0.585	0.569	-1.045	-0.517
Canada	-3.986**	0.265	-0.566	-1.713*	-2.201**
Denmark	-2.762	-0.296	-0.328	-2.270**	-1.694*
Finland	-6.266**	-1.889*	-4.008**	-4.258**	-4.729**
France	-2.336	0.069	-0.933	-1.210	-1.154
Germany	-3.320*	0.084	-0.549	-1.869*	-2.592**
Italy	-2.165	1.471	-0.128	-1.782*	-1.437
Japan	-2.238	2.232	-0.360	-0.528	-0.791
Mexico	-6.331**	-4.629**	-3.925**	-4.597**	-4.089**
Netherlands	-1.646	1.656	0.300	0.280	0.386
Norway	-2.538	1.949	0.646	-1.744*	-1.175
Portugal	-2.893	0.438	-2.356**	-1.938*	-1.991**
Spain	-2.406	1.165	-1.342	-2.137**	-1.681*
Sweden	-3.045	0.589	-.305	-1.254	-1.768*
Switzerland	-1.419	1.640	0.644	-0.822	0.111
UK	-3.002	0.160	-1.334	-2.187**	-1.601

*significant at the 5% level **significant at the 1% level

$$\Delta f_t = d_{11} + \delta_1\left(f_{t-1} - \alpha - \beta p_{t-1}^{US}\right) + \phi \Delta p_t^{US} + \nu_t \qquad \text{Without instruments}$$

$$\Delta f_t = d_{11} + \delta_1 w_t + \phi \Delta p_t^{US} + \nu_t \qquad \text{With instrument}$$

$$w_t = \left(f_{t-1} - \alpha - \beta' p_{t-1}^{US}\right) - \left(f_{t-m-1} - \alpha - \beta' p_{t-m-1}^{US}\right) m = 2,4,7,9$$

cointegration can be rejected at the 5% level or better for eleven of the nineteen countries. With m = 9, there are ten instances in which the null is rejected at the 5% significance level or better. It is a bit disconcerting that the rejections of the null when m = 9 are not all the same countries as the rejections when m = 7. More specifically, the estimated δ_1 for Australia, Italy, Norway, Sweden and the UK are insignificant for either m = 7 or m = 9, but not both. These results suggest some sensitivity to the choice of m.

Breusch-Godfrey Lagrange multiplier (LM) tests for serial correlation are applied to the estimated error correction model and the various specifications with instrumental variables.[6] Results for fourteen of the countries in all the equations estimated using instruments show evidence of autocorrelations. The p values on the obs*R^2 statistics are all less than 10%. Given that an objective of this paper is to compare conclusions concerning PPP for the ECM and the ECM with instruments, rather than adding lags of the dependent variable Δf_t (and Δp_t^{US} in a few cases) to eliminate autocorrelation in each equation, the IV model with m = 9 is used to determine the lag specification. Specifically, if the marginal significance level for the Obs*R^2 stat is .15 or less up to 6 lags of the dependent variable are added to the IV estimation with m = 9 of the ECM to address autocorrelation. Lags are added until the marginal significance level exceeds .15. In three cases; Finland, Netherlands, Portugal; serial correlation persists even with 6 lags of Δf_t. In these cases one lag each of the dependent variable and Δp_t^{US} are added until the LM test produces a p value exceeding .15. Once a specification free of autocorrelation is obtained for the IV estimation with m = 9, the same lag structure is applied to all other equations estimated with instrumental variables and to the error correction model.[7] Table 2 shows the final decisions on lags and the resulting p-values on the LM tests obs*R^2 statistic using the m = 9 specification.

Results using the lag structure specified in Table 2 are displayed in Table 3. The estimated coefficients are shown in the row with the country name and their corresponding t-statistics are in the row immediately following. The correction for serial

[6] A table showing these results is available from the author. Four lags were used in the LM tests.
[7] Finland presents a problem because a specification without serial correlation using these criteria has not yet been found.

correlation changes the conclusions somewhat from those based on Table 1. Based on the ECM estimation with p values calculated using the Ericsson and MacKinnon

Table 2 Final Specification of the ECM and ECM with Instruments

Country	Lags Δf_t	Lags Δp_t^{US}	p-value
Argentina	4	0	.588
Australia	0	0	.185
Belgium	2	0	.353
Brazil	4	0	.326
Canada	0	0	.407
Denmark	0	0	.247
Finland	?	?	?
France	1	0	.214
Germany	0	0	.876
Italy	1	0	.166
Japan	1	0	.264
Mexico	1	0	.412
Netherlands	1	1	.696
Norway	4	0	.328
Portugal	2	2	.925
Spain	4	0	.155
Sweden	1	0	.250
Switzerland	2	0	.189
UK	0	0	.169

program, there is evidence of a cointegrating relation, hence support for PPP, in the data for Belgium, Canada, Germany, Japan, Mexico, and Sweden. When using the instruments, ILE suggest using larger values of m when lagged variables are introduced to address serial correlation. Thus, the IV estimation results with m = 2 are reported only for the no lag case and those for m = 4 only if zero or one lags are added. Furthermore, an additional IV estimation with m = 12 is included for each country. As can be seen from the table in the error correction model the null hypothesis of no cointegration can be rejected for 6 countries.[8]

[8] Again, the Ericsson and MacKinnon program is used to determine p values for the EC model.

Table 3 Estimated δ_1 and t-statistic for Equation (5) Without and With Instruments: Corrected for Autocorrelation

		Instrumental Variables				
Country	ECM	m = 2	m = 4	m = 7	m = 9	m = 12
Argentina	-0.326	NA	NA	0.224	0.006	-0.259
t stat	-2.481	NA	NA	0.810	0.027	-1.259
Australia	-0.086	0.139	-0.057	-0.093	-0.138*	-0.067
t stat	-2.026	0.986	-0.624	-1.164	-1.720	-0.817
Belgium	-0.401**	NA	NA	-0.209	-0.236	-0.274*
t stat	-4.275	NA	NA	-1.342	-1.515	-1.656
Brazil	-0.167	NA	NA	-0.180	-0.104	-0.134
t stat	-2.899	NA	NA	-1.466	-0.973	-1.429
Canada	-0.223**	0.0355	-0.052	-0.138*	-0.175**	-0.261**
t stat	-3.986	0.265	-0.566	-1.713	-2.201	-3.225
Denmark	-0.142	-0.034	-0.029	-0.177**	-0.134*	-0.088
t stat	-2.762	-0.296	-0.328	-2.270	-1.694	-1.113
Finland	?	?	?	?	?	?
t stat						
France	-0.181	NA	-0.267*	-0.225*	-0.248*	0.036
t stat	-3.023	NA	-1.800	-1.735	-1.825	0.206
Germany	-0.169*	0.011	-0.05	-0.137*	-0.184**	-0.168**
t stat	-3.320	0.083	-0.549	-1.869	-2.592	-2.321
Italy	-0.163	NA	-0.086	-0.188**	-0.153*	-0.039
t stat	-2.731	NA	-0.761	-2.165	-1.707	-0.425
Japan	-0.220**	NA	-0.282**	-0.143*	-0.166**	-0.166**
t stat	-4.560	NA	-3.313	-1.933	-2.363	-2.477
Mexico	-0.640**	NA	-0.550**	-0.617**	-0.566**	-0.634**
t stat	-6.802	NA	-3.812	-4.537	-4.013	-4.566
Netherlands	-0.100	NA	-0.109	-0.043	-0.022	-0.009
t stat	-2.474	NA	-1.156	-0.573	-0.309	-0.137
Norway	-0.117	NA	NA	-0.077	-0.072	-0.022
t stat	-2.077	NA	NA	-0.583	-0.604	-0.172
Portugal	-0.091	NA	NA	-0.141	-0.117	-0.078
t stat	-2.004	NA	NA	-1.465	-1.293	-0.830
Spain	-0.074	NA	NA	-0.110	-0.093	0.068
t stat	-1.629	NA	NA	-0.799	-0.772	0.574
Sweden	-0.268**	NA	-0.164	-0.177*	-0.218**	-0.146
t stat	-3.911	NA	-1.166	-1.750	-2.141	-1.359
Switzerland	-0.098	NA	NA	-0.059	-0.005	-0.002
t stat	-1.700	NA	NA	-0.626	-0.056	-0.028
UK	-0.154	0.023	-0.132	-0.183**	-0.131	-0.111
t stat	-3.002	0.160	-1.334	-2.187	-1.601	-1.416

NA-not reported due to the number of lags in the IV estimation.

For five of these six countries, the IV estimations with instruments constructed using both m = 7 and m = 9 also indicate rejection of the null, hence support for PPP. In the other instance, Belgium, the IV estimations do not indicate the presence of a cointegrating relation until using the instrument with m = 12. In two additional instances, France and Italy, the IV estimations with m = 7 and m = 9 indicate support for PPP although the error correction results do not. It is encouraging that the instrumental variable models for m = 7 and m = 9 lead to the same conclusions (unlike those reported earlier in Table 1) for all but two countries, Australia and the United Kingdom. In the later case, the t-statistic on $\hat{\delta}_1$ is significant (and negative) for m = 7 and just barely insignificant in the case of m = 9.

The ECM and the error correction model with instruments do not produce results as supportive of PPP as those from unit root tests reported by Taylor but they are broadly consistent with the findings of Lopez, Murray, and Papell and Wallace and Shelley. An interesting, but as of yet unexplained, observation is that support for PPP appears more likely for instruments with m larger than 2 or 4 but less than 12. Of course, this observation must be tempered by the fact that results are not reported for m = 2, 4 for a number of the countries due to the number of lags present in the specifications adjusted for serial correlation.

As shown earlier, the error correction model of equation (5) can be rewritten as the autoregressive distributed lag (ADL) model of equation (6). Preliminary results are provided initially in Table 4 with tests and adjustments for serial correlation reported subsequently. Generally, the initial findings are similar to those for the error correction model. Only four of the $\hat{\delta}_1$ estimated in the ADL version are significant at the 5% level or better, thus the null of no cointegration (PPP does not hold) cannot be rejected in 15

Table 4 t-Statistics from Estimation of Equation (6) Without and With Instruments

		Instrumental Variables			
Country	ADL MODEL	m = 2	m = 4	m = 7	m = 9
Argentina	-5.009**	-1.336	-3.886**	-1.674*	-2.867**
Australia	-1.878	0.712	-0.725	-1.159	-1.710*
Belgium	-3.216	-0.914	-2.222**	-1.833*	-1.873*
Brazil	-2.338	-0.483	0.451	-1.122	-0.586
Canada	-4.027*	-0.032	-0.827	-1.836*	-2.324**
Denmark	-2.680	-0.332	-0.357	-2.273**	-1.650*
Finland	-6.614**	-3.837**	-4.947**	-4.854**	-5.239**
France	-2.126	-0.241	-0.993	-1.192	-1.182
Germany	-3.329	0.543	-0.397	-1.896*	-2.576**
Italy	-2.019	1.121	-0.431	-1.822*	-1.480
Japan	-2.241	2.176	-0.411	-0.842	-1.057
Mexico	-6.333**	-4.799**	-3.841**	-4.445**	-4.004**
Netherlands	-1.490	1.075	0.047	0.235	0.357
Norway	-2.625	0.630	-.841	-2.284**	-1.706*
Portugal	-2.858	0.203	-1.069	-1.002	-1.209**
Spain	-2.379	0.108	-1.365	-1.579	-1.233
Sweden	-3.147	-1.000	-1.550	-2.018**	-2.580**
Switzerland	-1.310	0.819	0.013	-1.044	-0.103
UK	-2.937	-0.423	-1.644	-2.030**	-1.461

*significant at the 5% level **significant at the 1% level. For the ADL these are based on critical values reported in Banerjee, Dolado, and Mestre (1998).

$\Delta f_t = d'_{11} + \delta_1 f_{t-1} + \gamma p_{t-1}^{US} + \phi' \Delta p_t^{US} + \nu_t$ Without instruments

$\Delta f_t = d'_{11} + \delta_1 w_{1t} + \gamma w_{2t} + \phi' \Delta p_t^{US} + \nu_t$ With instruments

$w_t = (w_{1t}, w_{2t}) = \left[(f_{t-1} - f_{t-m-1}), (p_{t-1}^{US} - p_{t-m-1}^{US}) \right]$

---

cases. Using instruments reveals little support for PPP for smaller values of m. However, instruments having larger m (7 or 9) are more likely to produce rejections of the null, hence support for PPP. For eleven countries the IV estimations for m = 7 reveal support for PPP. In nine of these instances, the IV estimations with m = 9 also indicate rejection of the null. Estimations with instruments using m = 7 but not m = 9 indicate support for PPP in two additional countries (Italy and UK), while IV estimations with m = 9 but not m = 7 suggest PPP holds for Australia.

There are two notable differences from the ECM results of Table 1. The null of no cointegration is rejected for Belgium in the ADL model, but not the ECM, with instruments for m = 4, 7, 9 while the no cointegration null is rejected for the ECM version with instruments for m = 4,7,9, but not the ADL model, in the case of Portugal. Again, conclusions based on these results are by no means definitive. As with the ECM approach, serial correlation affects the ADL results.

The same basic procedure is followed in addressing serial correlation in the ADL model as used for the ECM version. LM tests with four lags are applied to the various specifications. In instances where there is evidence of serial correlation, up to 6 lags of the dependent variable are added to the IV estimation for m = 12 to eliminate the problem. In five situations, serial correlation did not disappear with lags of the dependent variable so an equal number of lags of Δf_t and Δp_t^{US} are added until an LM test produces an obs*R^2 statistic with a p value of .15 or better.[9] The final specification for each country along with the p value is shown in Table 5 and the estimated δ_1 and corresponding t-statistics are given in Table 6. As with the ECM, the purpose is to compare results for different values of m so the same number of lags is added to the ADL estimation and the versions estimated with instruments.

[9] A table reporting the LM test results is available from the author.

Table 5 Lags Included to Eliminate Autocorrelation

Country	Lags of Δf_t	Lags of	p value of LM test with 4
Argentina	0	0	.338
Australia	0	0	.395
Belgium	1	0	.186
Brazil	4	0	.377
Canada	3	0	.170
Denmark	0	0	.337
Finland	1	0	.396
France	4	0	.232
Germany	0	0	.916
Italy	2	0	.208
Japan	1	0	.270
Mexico	1	0	.669
Netherlands	1	1	.391
Norway	4	0	.186
Portugal	2	2	.833
Spain	2	2	.629
Sweden	1	1	.764
Switzerland	1	1	.381
UK	0	0	.226

Table 6 Estimated δ_I and t-statistic for Equation (6) Without and With Instruments: Corrected for Autocorrelation

		Instruments				
Country	ADL	m = 2	m = 4	m = 7	m = 9	m = 12
Argentina	-0.396	-0.208	-0.448**	-0.198*	-0.290**	-0.431**
t-stat	-5.009	-1.336	-3.886	-1.674	-2.867	-4.032
Australia	-0.080	0.108	-0.069	-0.094	-0.138*	-0.068
t-stat	-1.878	0.712	-0.725	-1.159	-1.701	-.813
Belgium	-0.494	NA	-0.520**	-0.361**	-0.379**	-0.436**
t-stat	-5.346	NA	-3.450	-2.686	-2.667	-2.576
Brazil	-0.163	NA	NA	-0.229*	-0.128	-0.172*
t-stat	-2.779	NA	NA	-1.805	-1.188	-1.843
Canada	-0.217	NA	NA	-0.125	-0.186**	-0.254**
t-stat	-3.297	NA	NA	-1.317	-2.069	-2.731
Denmark	-0.138	-0.038	-0.031	-0.178**	-0.133*	-0.087
t-stat	-2.680	-0.332	-0.357	-2.273	-1.650	-1.092
Finland	-0.623	NA	-0.647**	-0.642**	-0.766**	-0.721**
t-stat	-7.979	NA	-5.716	-5.414	-5.826	-5.023
France	-0.115	NA	NA	-0.163	-0.195	0.168
t-stat	-1.652	NA	NA	-0.906	-1.154	0.707
Germany	-0.170	0.102	-0.035	-0.137*	-0.180**	-0.167**
t-stat	-3.329	0.543	-0.397	-1.896	-2.576	-2.309
Italy	-0.190	NA	NA	-0.250**	-0.187*	-0.066
t-stat	-2.957	NA	NA	-2.544	-1.912	-0.662
Japan	-0.226	NA	-0.273**	-0.156**	-0.177**	-0.189**
t-stat	-4.650	NA	-3.532	-2.277	-2.621	-2.825
Mexico	-0.638	NA	-0.555**	-0.624**	-0.567**	-0.636**
t-stat	-6.767	NA	-3.673	-4.321	-3.887	-4.510
Netherlands	-0.089	NA	-0.107	-0.039	-0.017	-0.001
t-stat	-2.194	NA	-1.109	-0.517	-0.236	-0.016
Norway	NA	NA	NA	-0.211	-0.137	-0.084
t-stat	NA	NA	NA	-1.516	-1.074	-0.595
Portugal	-0.072	NA	-0.193	-0.155*	-0.142	-0.095
t-stat	-1.553	NA	-1.455	-1.652	-1.613	-1.031
Spain	-0.084	NA	-0.201	-0.186*	-0.123	0.032
t-stat	-1.980	NA	-1.293	-1.728	-1.208	0.326
Sweden	-0.260	NA	-0.257**	-0.230**	-0.282**	-0.241*
t-stat	-3.599	NA	-2.042	-2.254	-2.623	-1.887
Switzerland	-0.141	NA	-0.179	-0.172**	-0.072	-0.089
t-stat	-2.485	NA	-1.556	-1.972	-0.811	-1.006
UK	-0.154	-0.070	-0.177	-0.178**	-0.122	-0.109
t-stat	-2.937	-0.423	-1.644	-2.030	-1.461	-1.366

*significant at the 5% level **significant at the 1% level.

As can be seen in the table, when equation (6) is estimated with instruments the null of $\delta_1 = 0$ can be rejected in at least 3 of the 4 IV specifications (m = 4,7,9,12) in seven instances. For four additional countries two of the IV estimations produce estimated values of δ_1 that are significantly negative hence supportive of PPP. In four more cases the null is rejected in one of the equations estimated with instruments. As with the ECM results, support for purchasing power parity is most often seen in estimations using larger values of m. Indeed, the null of no cointegration is rejected for fourteen countries with instruments constructed using m = 7 and in eleven instances for m = 9. On the one hand, it is once again disconcerting to find results so dependent on the value of m. More positively, the different results may simply suggest that the test with instruments has low power when a suboptimal instrument is used. Clearly additional work is needed to establish criteria for selecting m.

Finally, Table 7 shows the t-statistics for the estimated δ_1 from the second step of the Engle-Granger (EG) procedure equation (7) compared to those derived from the EG approach with instruments replacing the estimated residuals, $\left(f_{t-1} - \hat{\alpha} - \hat{\beta} p_{t-1}^{US}\right)$. The null hypothesis is $\delta_1 = 0$. Failure to reject the null indicates the presence of a unit root in the estimated equation, that is, the absence of a cointegrating relation between the dollar-denominated foreign price level and the US price level over the sample period. In other words, failure to reject the null would signal failure to support PPP. As results in Table 7 show, unit root tests applied to the estimated errors residuals from equation (7) for each country clearly reject the null in all instances.[10] The null is rejected, thus PPP is supported, at the 5% level for nearly as many countries (eighteen) using an instrument constructed with m = 7. The results for the instrument with m = 9 are similar.

[10] Akaike information criterion (AIC) is used to determine the number of lags in the unit root test. The number of lags chosen by AIC is then imposed on the estimations with instruments.

Table 7 t-Statistics on Estimated δ_1 for Equation 7 Without and With Instruments

			Instruments		
Country	Lags	Unit Root Test	m = 4	m = 7	m = 9
Argentina	0	-4.853**	-4.347**	-1.772*	-2.849**
Australia	1	-3.041**	-1.902*	-2.106**	-2.796**
Belgium	1	-5.519**	-3.421**	-3.077**	-3.141**
Brazil	4	-3.059**	NA	-1.922*	-1.318
Canada	2	-4.300**	-1.037	-2.052**	-2.801**
Denmark	1	-3.790**	-1.102	-3.062**	-2.497**
Finland	1	-6.266**	-4.333**	-4.298**	-4.578**
France	2	-3.177**	-1.683*	-2.333**	-2.466**
Germany	1	-3.564**	-0.716	-1.912*	-2.242**
Italy	2	-4.267**	-2.162**	-2.995**	-2.612**
Japan	1	-5.290**	-4.040**	-2.888**	-3.075**
Mexico	1	-7.018**	-4.000**	-4.565**	-3.943**
Netherlands	1	-3.575**	-1.949*	-1.347	-1.494
Norway	1	-4.147**	-2.108**	-3.054**	-2.482**
Portugal	5	-2.155*	NA	-1.812*	-1.288
Spain	1	-3.212**	-2.742**	-2.848**	-2.288**
Sweden	1	-4.441**	-2.027**	-2.434**	-2.824**
Switzerland	1	-4.104**	-2.163**	-2.667**	-1.891*
UK	0	-3.191**	-1.545	-2.359**	-1.787*

NA-not reported due to the large number of lags in the estimation.

Conclusions

These findings certainly do not resolve questions about purchasing power parity. The ECM and ADL model, especially when using instrumental variables, provide some support for the PPP hypothesis while the EG procedure with or without instruments produces strong evidence in favor of purchasing power parity. Indeed the EG results without instruments state that PPP with the US dollar holds for every country in the sample. The EG results with instruments, especially for m = 7,9 are nearly as strong.

How does the ILE instrumental variable test for cointegration compare? The ILE approach certainly simplifies single equation cointegration tests in that the asymptotic properties of the t statistics are standard normal. But, as with any newly developed procedure there are many questions remaining to be answered. Why do results, sometimes, differ across different values of m? Why are rejections of the null more likely for larger m, at least with the Taylor data? What is the optimal value of m? Why do conclusions regarding purchasing power parity differ so much across tests? Hopefully, additional research will shed light on these issues.

References

Banerjee, A; J.J. Dolado; and R. Mestre, 1998. Error-correction mechanism for cointegration in a single-equation framework, *Journal of Time Series Analysis* 19, 268-283.

Elliott, G., T. Rothenberg, and J.H. Stock, 1996. Efficient tests for an autoregressive unit root, *Econometrica* 64, 813-836.

Enders, W.; J. Lee; and M.C. Strazicich, 2007. New threshold cointegration tests and the Taylor rule. Unpublished working paper.

Enders, W. *Applied Econometric Time Series*, 2nd edition, John Wiley & Sons, Hoboken, NJ, 2004.

Ericsson, N.R. and J.G. MacKinnon, 2002. Distributions of error corrections tests for cointegration, *Econometrics Journal* 5, 285-318. Computer program available at http://www.econ.queensu.ca/faculty/mackinnon/.

Im, K.S; J. Lee; and W. Enders, 2006. Cointegration tests using instrumental variables with an example of the U.K. demand for money. Unpublished working paper.

Kremers, J.J.M.; N.R. Ericsson; and J.J.Dolado, 1992. The power of cointegration tests, *Oxford Bulletin of Economics and Statistics* 54, 325-348.

Kwiatkowski, D.; P.C.B Phillips; P. Schmidt; and Y. Shin, 1992. Testing the null hypothesis of stationarity against the alternative of a unit-root, *Journal of Econometrics* 54, 159-178.

Lopez, C.; C.J. Murray; and D.H. Papell, 2005. State of the art unit root tests and purchasing power parity, *Journal of Money, Credit, and Banking,* 37 (April) (2) 361-369.

Pesavento, E., 2007. Residuals based tests for the null of no cointegration: An analytical comparison, *Journal of Time Series Analysis,* 28 (January) (1) 111-137.

Rogoff, K., 1996. The purchasing power parity puzzle. *Journal of Economic Literature* 34, 647-668.

Sarno, L. and M.P. Taylor, 2002. The Economics of Exchange Rates. Cambridge University Press, Cambridge.

Taylor, A.T., 2002. A century of purchasing-power parity. *Review of Economics and Statistics* 84, 139-150.

_________ and Taylor M.P., 2004. The purchasing power parity debate. *Journal of Economic Perspectives* 18, 135-158.

Wallace, F.H. and G.L. Shelley, 2006. An alternative test of purchasing power parity, *Economics Letters,* 92 (August) (2) 177-183.