A long memory panel unit root test: PPP revisited.

Jonas Andersson* and Johan Lyhagen†

Working Paper Series in Economics and Finance, No. 303
Stockholm School of Economics

February 12, 1999

Abstract

In this paper two techniques, long memory and panel data models, are combined in order to increase the power of unit root tests. The power is shown to be always better against fractional alternatives and usually against autoregressive alternatives. The test is then used to reanalyze data sets investigated by Cheung and Lai (1993), Oh (1996) and Papell (1997) who studied the purchasing power parity. In some cases, the test rejected the hypothesis of no cointegration between foreign and domestic prices where the other authors tests did not.

Keywords: Unit root; Long memory; Purchasing power parity

JEL: C12; C14; C22; C23

*Department of Statistics, Uppsala University, P.O. Box 513, SE-751 20 Uppsala, Sweden. E-mail: jonas.andersson@statistik.uu.se

†Department of Economic Statistics, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm, Sweden. E-mail: stjl@hhs.se. Financial support from the Tore Browaldhs Foundation for Scientific Research and Education is greatfully acknowledged.

1 Introduction

Dynamic models is an important class of models for modeling economic relations. Before one begins to describe the underlying process with a model, some properties of the included variables must be established. In the case of dynamic models, the existence of unit roots is probably the most important one, a view which is supported by the large amount of literature published in this area. Even if the importance of unit roots was well established it was first when Dickey and Fuller (1979) provided the most used unit root test so far, that testing for unit roots become standard. Later, they extended their test in order to adjust for the effects of lags of Δx_t and deterministic trends (Dickey and Fuller, 1981). This test is commonly abbreviated as ADF. A nice treatment of unit roots is in Campbell and Perron (1991).

Most often economic problems, that may be modeled in a dynamic framework, is analyzed in the context of one item, where an item often is a country. In other areas of economics the use of panels of data are often used, e.g. unemployment analysis or other micro economic areas. Although the introduction of dynamic panel models, Balestra and Nerlove (1966), are of some age it is only recently that unit root testing in panels has been considered. Quah (1994) is by our knowledge the first published work and a few applied papers have been published after that, e.g. McDonald (1996), Oh (1996), Coakley and Fuertes (1997), Culver and Papell (1997) and Papell (1997).

Recent studies, (e.g., Diebold and Rudebusch, 1991; Cheung and Lai, 1993) indicates that a test based on a long memory model instead of a test of Dickey-Fuller type gives higher power against long memory alternatives. Surprisingly, this is sometimes also true when the data generating process is an AR process. In this paper we will follow up this idea to increase the power compared to ADF. More precisely, a panel form of the GPH test, see Geweke and Porter-Hudak (1983), is proposed, abbreviated as PGPH, and compared to the panel ADF (PADF) used e.g. by Culver and Papell (1997) and Papell (1997).

We investigate the purchasing power parity (PPP) relation in an empirical part of this paper. The cases we investigate is when other authors have failed to reject the null of a stationary relationship. The data is annual, quarterly and monthly previously used by Cheung and Lai (1993), Oh (1996) and Papell (1997).

In the next section, the panel GPH test is presented and in Section 3 the power of it is compared with the panel ADF test. Section 4 presents a revision of the tests of the PPP in Cheung and Lai (1993), Oh (1996) and Papell (1997) where the new test is used. A conclusion closes the paper.

2 The long memory panel data test

The power of the ADF test has been shown to be rather low against long memory alternatives, see Cheung and Lai (1993). This kind of memory is interesting because of at least two reasons. Firstly, it is possible to test unit roots against mean reversion instead of integration of order zero. This means that the alternative in some sense is broader than in the ADF test. Secondly, if a long memory model is used the large number of parameters can be avoided, at the cost of a subjective choice of one truncation parameter.

The panel ADF-test is based on the regression

$$\Delta x_{jt} = \mu_j + \alpha x_{jt-1} + \sum_{i=1}^{k} c_i \Delta x_{t-i} + \varepsilon_t \tag{1}$$

and the hypothesis of a unit root is equivalent to $\alpha = 0$. The lags of Δx_t is present to whitening the residuals and is not of importance by themselves. In order to decide the number of lags, k, Campbell and Perron (1991) proposed a recursive data dependent method. By choosing a maximum value, $k_{\max}$, estimating the model and then remove the last lag if its regression coefficient is not significant, k can be determined.

A panel unit root test can be constructed by estimating the equations $j = 1, 2, ..., N$ where N is the number of time series. It can be estimated by GLS with the restriction that α is equal for all j as done by Culver and Papell (1997) and Papell (1997). The null hypothesis is that all time series contains a unit root and the null distribution can be obtained by simulation.

The long memory approach to the unit root problem was proposed by Cheung and Lai (1993). Their test is based on the estimator of long memory proposed by Geweke and Porter-Hudak (1983). The appeal of this method is that it can estimate the long memory parameter without specifying the short term structure of the process. Instead a truncation parameter α need to be chosen to decide how many frequencies, m, that are going to be used in the regression, i.e. $m = T^{\alpha}$. This choice of m will give consistency of the estimator, if additionally the smallest frequencies are excluded normality is gained (e.g., Geweke and Porter-Hudak, 1983; Robinson, 1995). A consequence is that we avoid the data dependent procedure to determine the number of lags that we have to do using the ADF test at the cost of one truncation parameter. The GPH estimator is based on the regression of the m lowest frequencies in

$$\ln I(\omega) = C - 2d \ln \left|1 - e^{-i\omega}\right| + \ln \xi_\omega \tag{2}$$

The unit root test is made by using the $t-$statistic on differenced data of the hypothesis $d = 0$ against $d < 0$. The proposed panel data test is simply the mean of the individual $t-$test.

Often the T^α smallest frequencies is used where α is set to $0.5 - 0.6$ but Andersson and Lyhagen (1997) show that choosing an α greater then 0.6 improves the power substantially in the case when $N = 1$. Note that the fully parametric ARFIMA(p, d, q) model may be used in a panel unit root test instead for the ADF regressions.

3 Power comparison

The null process is a pure random walk with no short or long memory, i.e.

$$(1 - L)\, x_t = e_t.$$

The sample sizes considered are $T = 50$ and $T = 100$ and the number of replicates is 50000. We use the PADF test which implies that a procedure of choosing the lag order must be decided upon. A maximum lag order of 10 is decided and a general to specific procedure is carried out by the use of a $t-$test. The critical values for tests on the 5%-level of the PADF test statistic and the PGPH test statistic are shown in Table 1.

Table 1-5 *in here*

In the power comparison two data generating processes are considered under the alternative hypothesis. The first is the AR(1) process and the second is the fractional noise process, which is a special case of the ARFIMA class of processes, see Granger and Joyeux (1980) or Hosking (1981). The short memory parameter and the long memory parameter are each varied from 0.1 to 0.9 in steps of 0.1. The results of the Monte Carlo simulation are shown in Tables 2-5. The panel GPH test had always better power against fractional alternatives. The same was true also against the AR(1) alternatives, except when the autoregressive parameter α was 0.9. The difference in power decrease as N increases for both alternatives. The reason for this is that the asymptotic power is one and that this limit is gained for small values of N. The difference may still be large, e.g. for $N = 2$ and $d = 0.5$ the power for PADF is around 14% but 96% for the PGPH ($T = 50$). The power figures for the AR alternative with the same value ($\alpha = 0.5$) is the same for PGPH but 39% for PADF. A conclusion from this result is that PADF is much more sensitive against which alternative that generates data than PGPH. A small increase of T from 50 to 100 may have a substantial effect on the power, e.g. for $N = 2$ and $\alpha = 0.1$, a very distant alternative, the power increases from 49% to 94%. for the PADF.

4 PPP revisited

In this part of the paper we investigate some cases where other authors have failed to reject the null of unit root/no cointegration in PPP data. The PPP relationship is simply stated in Cheung and Lai (1993) as

$$sp_t = \alpha_0 + \alpha_1 p_t + e_t$$

where α_0 is some constant, sp_t is the foreign price index converted to domestic currency units, p_t is the domestic price index, and e_t is an error term. All variables are in logarithms. For the PPP to hold e_t should be mean-reverting (see Cheung and Lai (1993) for a formal definition of a mean-reverting process). In Oh (1996) and in Papell (1997) α_0 and α_1 are assumed to take their theoretical values (0 and 1). Cheung and Lai (1993) fixed the augmentation to 4 lags, Oh (1996) uses lags 1 to 4 while Papell (1997) uses the ones gained in the univariate analysis.

The first paper of the three considered is Cheung and Lai (1993) who use the GPH test as a test for (fractional) cointegration. The data is annual for the period 1914-1989 for United States, United Kingdom, France, Italy, Canada, and Japan and the US-dollar is the base currency. The data is collected from Lee (1976) and the International Monetary Fund's International Financial Statistics data tape. The null is only rejected for Italy. Here, the panel GPH test will be used on the five country group.

Oh (1996) uses yearly data between 1960 and 1989 taken from the Penn World Tables (PWT), Mark 5.5. The data is analyzed both for all the 111 countries and separately for development and OECD countries respectively. For a longer period starting 1950 a group of 51 countries are studied. This time period is also divided into sub periods. Finally, the tests are applied to the G-6[1] countries. The null is not rejected when the OECD and G-6 are analyzed for the periods 1950-1972 and 1950-1990. These are the cases for which the PGPH is to be used.

Instead of annual data, as in the two papers above, Papell (1997) uses quarterly and monthly data. It is taken from the International Monetary Fund's International Financial Statistics data tape. The intention was to investigate PPP for OECD countries but due to lack of observations and other considerations a reduced number of countries was used. Luxembourg has a currency union with Belgium and Iceland has lack of data so for quarterly data the number of countries is 21. Three countries had not monthly data, Australia, Ireland and New Zealand, which gives 18 countries. The series starts in January 1973 and ends in September 1994, i.e. 87 quarterly and 261 monthly observations. When using US as base currency and quarterly data the unit root hypothesis is never rejected and for monthly data the EMS and the G-6 do not reject the null hypothesis. For the German base the result is similar except that the All 20 group rejects the null, and for monthly data the result is the same as when using US as base. For these non-rejecting cases the PGPH test will be used.

The differences between their results and the results we obtained by using the PGPH test can be seen in Table 6. The null hypothesis of no cointegration is rejected for the five countries group in Cheung and Lai (1993). In their study, the null is only rejected for one of them, Italy. For the data used by Oh (1996), the null hypothesis is not rejected when we use the PGPH test. The null is only rejected for the EC-countries, when Germany is the basis currency, for the data considered in Papell (1997) .

Table 6 *in here*

5 Conclusions

A long memory panel unit root test has been suggested and compared to the panel Dickey-Fuller test. The power is larger against both fractional and first order autoregressive alternatives except for the autoregressive case with parameter 0.9, the value closest to the null hypothesis.

[1] The G6 countries are Canada, France, Germany, Italy, Japan and United Kingdom.

The test was then used to reanalyze some previous empirical studies of the power purchase parity by Cheung and Lai (1993), Oh (1996) and Papell (1997). The general conclusion in these papers was that the null hypothesis of no cointegration between foreign and domestic prices could seldom be rejected. Two differences between these results and ours was observed in the study. The null hypothesis was rejected by the PGPH test for the panel of countries in Cheung and Lai (1993). In their study, it was just for one of the five countries, Italy, that it was rejected. The other difference was our rejection for Germany when we applied the PGPH test to Papells data. One possible reason for the difference in results between the data from Cheung and Lai (1993) and the other is that Cheung and Lai (1993) estimated the cointegrating relation while Oh (1996) and Papell (1997) did not.

Even though the differences to the other studies are not striking, we believe that a long memory approach can be useful since it allows a slower convergence of the prices.

References

Andersson, J. and Lyhagen, J. (1997) A note on the power of the GPH test for cointegration *Research Report 97:7, Department of Statistics, Uppsala University.*

Balestra, P. and Nerlove, M. (1966) Pooling cross-section and time-series data in the estimation of a dynamic model: the demand for natural gas. *Econometrica* 34, 585-612.

Campbell, J.Y. and Perron, P. (1991) Pitfalls and opporunities: what macroeconomists should know about unit roots. *National Bureau of Economic Research Macroeconomics Annual* 141-201.

Cheung, Y.W. and Lai, K. (1993) A fractional cointegration analysis of purchasing power parity. *Journal of Business & Economic Statistics* 11, 103-112.

Coakley, J. and Fuertes, A.M. (1997) New panel unit root tests of PPP. *Economics Letters* 57, 17-22.

Culver, S.E. and Papell, D.H. (1997) Is there a unit root in the inflation rate? Evidence from sequential break and panel data model.

Journal of Applied Econometrics 35, 155-160.

DICKEY, D.A. and FULLER, W.A. (1979) Distribution of the estimator for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.

DICKEY, D.A. and FULLER, W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49, 1057-1072.

DIEBOLD, F. and RUDEBUSCH, G. (1991) On the power of Dickey-Fuller tests against fractional alternatives. *Economics Letters* 35, 155-160.

GEWEKE, J. and PORTER-HUDAK, S. (1983) The estimation and aplication of long memory time series models. *Journal of Time Series Analysis* 4, 221-238.

GRANGER, C.W. and JOYEUX, R. (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 15-29.

HOSKING, J.R.M. (1981) Fractional differencing. *BiometriKa* 68, 165-176.

MCDONALD, R. (1996) Panel unit root tests and real exchange rates. *Economics Letters* 50, 7-11.

LEE, M.K. (1978) *Purchasing power parity* Marcel Dekker, New York.

OH, K-Y. (1996) Purchasing power parity and unit roots tests using panel datat. *Journal of International Money and Finance* 15, 405-418.

PAPELL, D. (1997) Searching for stationarity: Purchashing power parity under the current float. *Journal of International Economics* 43, 313-323.

QUAH, D. (1994) Exploiting cross sectional variation for unit root inference in dynamic data. *Economics Letters* 44, 9-19.

ROBINSON, P.M. (1995) Log-periodogram regression for time series with long-range dependence. *The Annals of Statistics* 23, 1048-1072.

Table 1: Critical values. Number of observations is T=50 and T=100, number of replicates is 50000.

$N=$	1		2	
T	$PADF$	$PGPH$	$PADF$	$PGPH$
50	-3.5452	-1.9717	-4.0557	-1.3281
100	-3.1113	-1.8102	-3.6148	-1.2588

$N=$	5		10		20	
T	$PADF$	$PGPH$	$PADF$	$PGPH$	$PADF$	$PGPH$
50	-5.2468	-.8036	-6.8078	-.5630	-9.7419	-.3928
100	-4.6615	-.7774	-5.9540	-.5419	-7.9898	-.3833

Table 2: Size adjusted power when the alternative is fractional noise. Number of observations is T=50, number of replicates is 10000.

$N =$	1		2	
d	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.2546	.9782	.4079	.9999
0.2	.2013	.9614	.3185	.9991
0.3	.1626	.9263	.2489	.9972
0.4	.1296	.8655	.2489	.9882
0.5	.1008	.7573	.1382	.9589
0.6	.0859	.5987	.1079	.8671
0.7	.0719	.4174	.0794	.6713
0.8	.0619	.2431	.0674	.4031
0.9	.0556	.1214	.0552	.1672

$N =$	5		10		20	
d	$PADF$	$PGPH$	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.7344	1.0000	.9357	1.0000	.9945	1.0000
0.2	.6081	1.0000	.8598	1.0000	.9764	1.0000
0.3	.4853	1.0000	.7528	1.0000	.9239	1.0000
0.4	.3602	1.0000	.5933	1.0000	.7919	1.0000
0.5	.2505	.9997	.4271	1.0000	.6093	1.0000
0.6	.1684	.9979	.2654	1.0000	.4021	1.0000
0.7	.1209	.9604	.1715	.9990	.2304	1.0000
0.8	.0831	.7405	.1063	.9512	.1388	.9985
0.9	.0621	.3019	.0698	.4944	.0805	.7581

Table 3: Size adjusted power when the alternative is fractional noise. Number of observations is T=100, number of replicates is 10000.

$N =$	1		2	
d	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.5782	1.0000	.8907	1.0000
0.2	.5097	.9999	.8194	1.0000
0.3	.4667	.9993	.7365	1.0000
0.4	.3880	.9971	.6390	1.0000
0.5	.3078	.9854	.4974	1.0000
0.6	.1874	.9301	.3114	.9983
0.7	.1127	.7578	.1746	.9631
0.8	.0787	.4700	.0990	.7293
0.9	.0671	.1983	.0735	.3030

$N =$	5		10		20	
d	$PADF$	$PGPH$	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.9998	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	.9968	1.0000	1.0000	1.0000	1.0000	1.0000
0.3	.9842	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	.9492	1.0000	.9984	1.0000	1.0000	1.0000
0.5	.8478	1.0000	.9848	1.0000	.9998	1.0000
0.6	.6309	1.0000	.8954	1.0000	.9925	1.0000
0.7	.3542	1.0000	.5859	1.0000	.8531	1.0000
0.8	.1655	.9782	.2572	1.0000	.4324	1.0000
0.9	.0832	.5714	.1080	.8322	.1487	.9808

Table 4: Size adjusted power when the alternative is an AR(1) process. Number of observations is T=50, number of replicates is 10000.

$N =$	1		2	
α	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.3171	.9755	.4905	.9998
0.2	.3008	.9577	.4656	.9985
0.3	.2834	.9229	.4500	.9965
0.4	.2784	.8542	.4242	.9867
0.5	.2557	.7343	.3872	.9501
0.6	.2303	.5700	.3408	.8341
0.7	.1919	.3741	.2759	.5972
0.8	.1338	.1895	.1953	.3196
0.9	.0920	.0894	.1152	.1125

$N =$	5		10		20	
α	$PADF$	$PGPH$	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.8163	1.0000	.9734	1.0000	.9991	1.0000
0.2	.8089	1.0000	.9686	1.0000	.9989	1.0000
0.3	.7869	1.0000	.9614	1.0000	.9979	1.0000
0.4	.7492	1.0000	.9490	1.0000	.9966	1.0000
0.5	.7142	.9998	.9200	1.0000	.9916	1.0000
0.6	.6388	.9956	.8821	1.0000	.9831	1.0000
0.7	.5433	.9331	.8065	.9968	.9475	1.0000
0.8	.3845	.6081	.6334	.8654	.8283	.9883
0.9	.1969	.1805	.3308	.2757	.4662	.4433

Table 5: Size adjusted power when the alternative is an AR(1) process. Number of observations is T=100, number of replicates is 10000.

$N =$	1		2	
α	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	.6548	1.0000	.9380	1.0000
0.2	.6434	.9998	.9300	1.0000
0.3	.6137	.9996	.9204	1.0000
0.4	.5978	.9969	.8965	1.0000
0.5	.5614	.9880	.8763	.9998
0.6	.5207	.9495	.8386	.9984
0.7	.4512	.8130	.7570	.9776
0.8	.3535	.5035	.6149	.7716
0.9	.1926	.1793	.3304	.2712

$N =$	5		10		20	
α	$PADF$	$PGPH$	$PADF$	$PGPH$	$PADF$	$PGPH$
0.1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.3	.9998	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	.9995	1.0000	1.0000	1.0000	1.0000	1.0000
0.5	.9996	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	.9973	1.0000	1.0000	1.0000	1.0000	1.0000
0.7	.9931	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	.9578	.9885	.9990	1.0000	1.0000	1.0000
0.9	.7000	.5218	.9434	.7983	.9982	.9702

Table 6: The panel GPH test applied to the data investigated in Cheung and Lai (1993), Oh (1996) and Papell (1997).

Article	Group	Base	Period	N	T	Freq.	PGPH	Crit.
C&L (1993)	All five	US	1914-1989	5	76	Y	-3.3965	-.7707
Oh (1996)	OECD	US	1950-1990	22	41	Y	1.2813	-.2619
Oh (1996)	OECD	US	1950-1972	22	23	Y	1.1215	-.2319
Oh (1996)	G6	US	1950-1990	6	41	Y	0.9959	-.6091
Oh (1996)	G6	US	1950-1972	6	23	Y	1.2987	-.6017
Papell (1997)	All 20	US	1973-1994	20	87	Q	0.4719	-.3088
Papell (1997)	EC[2]	US	1973-1994	11	87	Q	0.02344	-.4420
Papell (1997)	EMS[3]	US	1973-1994	7	87	Q	0.5260	-.5741
Papell (1997)	EMS	US	1973-1994	6	261	M	1.6212	-.6473
Papell (1997)	G6	US	1973-1994	6	87	Q	0.6195	-.6269
Papell (1997)	G6	US	1973-1994	6	261	M	1.4469	-.6473
Papell (1997)	G10[4]	US	1973-1994	10	87	Q	0.8398	-.4654
Papell (1997)	EC	Germany	1973-1994	10	87	Q	-0.7551	-.4654
Papell (1997)	EMS	Germany	1973-1994	6	87	Q	-0.1885	-.6269
Papell (1997)	EMS	Germany	1973-1994	5	261	M	0.5429	-.7135
Papell (1997)	G6	Germany	1973-1994	5	87	Q	0.0209	-.6922
Papell (1997)	G6	Germany	1973-1994	5	261	M	0.7784	-.7135
Papell (1997)	G10	Germany	1973-1994	9	87	Q	0.1817	-.4980

[2]EC consists of Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain and the United Kingdom.

[3]EMS consists of Belgium, Denmark, France, Germany, Ireland, Italy and the Netherlands.

[4]The G10 group consists of Belgium, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland and the United Kingdom.