Working Paper

INTERNATIONAL MONETARY FUND

IMF Working Paper

WP/99/50

INTERNATIONAL MONETARY FUND

Monetary and Exchange Affairs Department

Long-Run Exchange Rate Dynamics: A Panel Data Study

Prepared by Karl F. Habermeier and Mario Mesquita[1]

Authorized for distribution by Elizabeth Milne

April 1999

Abstract

Long-run movements of real exchange rates are studied using a panel data set comprising 51 economies. The purchasing power parity hypothesis (PPP) is examined first using unit root tests. It is found that PPP does not hold for the full sample of countries, but it may hold for the advanced economies, as well as open and high-inflation economies. Using the recently developed mean group and pooled mean group estimators, the paper finds support for the Balassa-Samuelson hypothesis in both advanced and developing economies; and for the influence of shifts in the terms of trade.

JEL Classification Numbers: F31

Keywords: Exchange Rates, Panel Data Econometrics, Purchasing Power Parity (PPP), Balassa-Samuelson Effect, Terms of Trade.

Authors' E-Mail Addresses: khabermeier@imf.org and mmesquita@imf.org

[1]The authors would like to thank their colleagues in the International Monetary Fund for many helpful comments. Any remaining errors are the authors'.

Contents Page

I. Introduction

The paper investigates some traditional propositions in exchange rate economics through an empirical multi-country analysis. A panel of 51 advanced and developing economies covering the floating exchange rate period (1971-1997) is used to test several hypotheses about the long-run behavior of real and nominal effective exchange rates. The paper re-examines earlier results obtained for some of the advanced economies, and offers new insights into the behavior of real exchange rates in developing economies.

To begin with, the purchasing power parity (PPP) hypothesis is re-examined using panel unit root tests. There are several different versions of the PPP hypothesis; we focus here on "long-run PPP", which asserts that the real exchange rate (i.e. the nominal exchange rate adjusted for relative prices) reverts to a constant long-run average. Long-run PPP rests on two key assumptions: the law of one price and long-run money neutrality. These assumptions are necessary but not sufficient: real shocks may require relative price adjustments which violate PPP. In statistical terms, long-run PPP means that real exchange rates should be stationary. The panel unit root tests reported on in this paper exploit cross-section variation to increase the power to reject the null hypothesis that real exchange rates are non-stationary, and hence to provide more robust support for PPP than single-equation tests conducted over a similar time span.

The results of the panel tests suggest that PPP does not hold for all countries, although it does appear to hold for the advanced economies, as well as for open and high-inflation economies. It is therefore interesting to investigate which other factors may account for persistent real exchange rate trends. The classic explanation for such trends is the Balassa-Samuelson effect, which under reasonable conditions implies that the real exchange rate of faster-growing economies would tend to appreciate. Additional factors that may help to account for long-run real exchange rate developments are permanent shifts in the terms of trade and measures of the size of the public sector. Recently developed panel data techniques, namely mean group and pooled mean group estimation, are used to find long-run parameters in several models of the real exchange rate. These estimates provide considerable support for the Balassa-Samuelson effect and for the influence of the terms of trade on real exchange rates. The role of the size of the public sector, as measured by the ratio of government consumption to GDP, is less clear.

Countries differ, and the results are found to depend to a large extent on the sub-sample of countries being examined. Countries are classified according to their stage of development, long-run average inflation rate, and openness to international trade. Moreover, panel data typically contain some degree of cross-section correlation, usually reflecting the presence of common factors, which may affect both the result of unit root tests and long-run parameter estimates. The tests were thus conducted both using the unadjusted data series and series from which the cross-sectional mean had been removed at each point in time.

The paper relies on real effective (trade weighted) exchange rates, which in principle are more likely to conform to PPP than simple bilateral real exchange rate indices, since effective rates are by construction more representative of countries' actual trade patterns. In addition, all of the other variables used are in relative or effective terms, e.g. the test of the Balassa-Samuelson hypothesis is based on a trade-weighted average of per capita real GDP.

Section II reports tests for PPP. An assessment of the Balassa-Samuelson, terms of trade, and government consumption effects is presented in Section III. Some concluding remarks are offered Section IV. Data sources and the formulae for construction of the data indices are presented in Appendices I and II.

II. Does PPP Hold?

Empirical testing for Purchasing Power Parity (PPP) has evolved considerably in the last thirty years. This evolution has been characterized by the continuous adoption of new econometric techniques and by a progressive dilution of the version of PPP being tested.[1]

Initially the focus was on *absolute PPP*, the notion that the equilibrium nominal exchange rate is such that the domestic and external purchasing powers of a currency are equal. Thus, under absolute PPP the real exchange rate is fixed and, in a strict interpretation, equal to unity. In the 1970s the focus shifted to *relative PPP*, which states that the equilibrium percentage rate of currency depreciation must equal the inflation differential between the home economy and the foreign economy, implying constancy of the real exchange rate. Since the 1980s, with the advent of the new time-series econometrics, researchers began to focus on the dynamic properties of real exchange rates. In particular, *long-run PPP* is deemed to hold when the real exchange rate is stationary (i.e. its distribution has finite and unvarying first and second moments). Most recently, there have been investigations of *cointegration PPP*, which requires only that the nominal exchange rate and domestic and foreign prices exhibit a stable relationship over time, but not one that necessarily satisfies the symmetry and proportionality restrictions implied by the traditional Casselian interpretation of PPP.

Statistical methods

Econometric techniques have changed from simple cross-section analysis, as in Balassa's (1964) classic work, to standard OLS based time-series, and more recently to integration and cointegration analysis, and to the introduction of persistence analysis, through impulse-response functions as well as variance ratio tests. Early evidence was not supportive of the PPP hypothesis. Absolute PPP was strongly rejected by Balassa's work, a result amply confirmed in more recent studies such as Heston and Summers (1991). Relative PPP also

---

[1]For surveys on PPP see Dornbusch (1988), Breuer (1994), Froot and Rogoff (1994), and Rogoff (1996).

found scant support, except for hyperinflationary economies.[2] Early tests of real exchange rate stationarity, which focused on developed countries in the recent floating exchange rate period, also tended to reject PPP. However, in a seminal work, Frankel (1986) showed that traditional, Dickey-Fuller-type unit root tests are likely to have low power to reject reasonable alternatives, such as slowly mean-reverting real exchange rates. Frankel's analysis cast considerable doubts on PPP tests based on small sample periods, and the search for higher power test specifications led to the use of long-period and multi-country tests (which add power by taking advantage of cross-sectional variation).[3]

The recent empirical literature has been much more supportive of PPP. The consensus now is that real exchange rates are indeed slowly mean-reverting, and that given a sufficiently long time span (fifty to seventy years), the hypothesis that they are nonstationary can often be rejected. In addition, tests based on cointegration analysis tend to confirm the existence of a stable long-run relation between nominal exchange rates and relative prices for several countries, although the specific relation supported by the data is often only distantly related to the theoretical priors. It also appears that PPP tends to hold for the core ERM countries, when the bilateral real exchange rates are in terms of the deutsche mark, more often than when bilateral rates are constructed with the US dollar as numeraire.

This literature has, however, focused on advanced countries, and evidence on the long-run behavior of real exchange rates, in particular PPP, in developing countries is still scarce.[4] Single-country studies, which require long time series, face clear data limitations in most developing countries.[5] An interesting alternative is to exploit the information content of cross-sectional variation to increase the power of the statistical tests. The best-known panel unit root tests have been developed by Levin and Lin (1993) and Im, Pesaran, and Shin (1997)—these tests will henceforth be abbreviated by LL and IPS. The economic interpretation of these tests is the same as in the case of individual country studies: rejection of the unit root hypothesis is evidence that real exchange rates revert to long-run means at a sufficiently rapid pace, as required by PPP.

Monte Carlo simulations in the IPS study show that the small sample properties of the IPS tests, in particular the t-bar test (to be described below), are superior to those of the LL test. The simulations covered several cases, notably serial correlation or the lack of it in the

---

[2]See, for instance, Krugman (1978) and Frenkel (1981).

[3]It should be noted that adding observations by using higher frequency data does not help, since the problem is one of not having sufficiently *long* samples.

[4]Edwards and Savastano (1998).

[5]There are however long-span studies for a few countries, for example Froot and Rogoff (1994) on Argentina, using the 1913-1988 sample and Mesquita (1996) for Brazil, using the 1870-1994 sample period.

underlying individual DF regressions, and the inclusion or exclusion of a time trend.[6] The experiments on ADF regressions without residual serial correlation indicated that for a sample with 25 periods of data for 50 groups (T=25 and N=50), the LL test has higher size than the t-bar test, but lower power.[7] The experiments on ADF regressions with serial correlation (with and without trend terms) showed that errors in the selection of the lag order may have dramatic effects on these tests. In fact, incorrectly using ADF(0) regressions when the correct lag order is higher tends to drive the size and power of these tests to zero. When the underlying regressions do not include trend terms, the IPS t-bar test has more power than the LL test. Moreover, the LL test is shown to suffer from significant size distortions; and the trade-off between size and power is much more favorable to the IPS t-bar test than to the LL test. When the underlying regressions include trend terms, lag length selection becomes even more critical. Again, the LL test suffers from significant size distortions, particularly as N rises in relation to T; and in general it has less power than the IPS t-bar test. Finally, O'Connell (1998) shows that correlation between the various country real exchange rate indices affects the size and power of the LL test. The IPS t-bar test can be made robust to correlation between groups by de-meaning the relevant series (across groups). Thus, it seems that on the whole the IPS t-bar test is more robust than the LL test, and it is therefore chosen for the empirical analysis that follows.

The IPS t-bar test averages the test statistics for the individual countries, and standardizing this average test statistic by its expected value and variance under the null hypothesis—which Im, Pesaran, and Shin have computed through simulations for various combinations of sample length, T, and ADF lag order. The resulting standardized test statistic, denoted Ψ-bar statistic, is distributed as a standard normal for large N, and its formula is:

$$\bar{\Psi} = \frac{\sqrt{N}\left(\bar{t}_{NT} - \frac{1}{N}\sum_{i=1}^{N} E[t_{iT}]\right)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} VAR[t_{iT}]}}$$

where $\bar{t}_{NT}$ is the average of the N individual ADF test statistics; and $E[t_{iT}]$ and $VAR[t_{iT}]$ are the empirical first and second moments of the ADF test statistics under the null. Calculation of the Ψ-bar statistic then involves running ADF regressions for the N countries, selecting the ADF regression lag length through some information criterion (here AIC was used), obtaining the cross-section average, and then constructing the test statistic shown above using the empirical expected values and variances in Im, Pesaran, and Shin (1997).

---

[6]A constant term was included in every case.

[7] The *power* of a test is the probability that the test statistic will reject a *false* null hypothesis, while the *size* of a test is the probability that the test statistic will reject a *true* null hypothesis.

Previous literature

Before describing our test results, some of the literature on PPP tests in panels is briefly reviewed.

MacDonald (1996) tests for PPP in OECD countries using annual data for 1973-92. Individual unit root tests confirm the established result that it is not possible to reject the unit root null for a vast majority of countries—it is rejected in only 5 out of 96 individual tests.[8] By contrast, the panel LL tests comfortably reject the null that all real exchange rates are nonstationary, regardless of the chosen deterministic specification and the price index used to construct real exchange rate indices.

Coakley and Fuertes (1997) use the IPS t-bar and LR-bar statistics to test for PPP in a sample of G-10 countries plus Switzerland, relying on monthly bilateral real exchange rates for the period from July 1973 to June 1996. The tests are carried out on real exchange rate series constructed from wholesale as well as consumer price indices. They conclude that while conventional ADF tests fail to reject the nonstationarity null, and Johansen tests for cointegration between nominal exchange rates and price indices yield mixed results, the IPS tests provide strong support for long-run PPP. Their IPS t-bar test results appear particularly robust for real exchange rates based on wholesale price indices, for which the unit root null is rejected at the 5 percent level for both the constant and constant plus trend specifications—consumer price based real exchange rates are only found to be stationary around a constant at the 10 percent level. LR tests confirm the results of the t-bar test.

Finally, Nagayasu (1998) conducts panel cointegration tests for PPP in a sample of 16 African countries, and finds support for a 'semi-strong' version of the hypothesis, according to which domestic and foreign prices affect the nominal exchange rate in a symmetric although not necessarily proportional way.

Test results

The tests presented below are based on a sample of the 51 largest market economies in the world, which jointly account for about 80 percent of world output, covering the period 1971-1997.[9] Real exchange rates are based on consumer price indices, as long series of wholesale price indices are often unavailable for developing countries, and were calculated in

---

[8]MacDonald (1996) finds that real exchange rate indices based on wholesale prices for Australia, Greece, and Spain are stationary around a constant, whereas consumer price based real exchange rate indices for Greece and Germany are stationary around a constant and a trend, respectively.

[9]Of the major market economies, only Hong-Kong had to be excluded for data limitations.

effective (i.e. trade-weighted) terms.[10] The sample is substantially wider than in previous panel studies, and includes 26 developing and 25 advanced countries.[11]

Table 1 below shows the results of individual-country ADF tests under two specifications for the deterministic terms. It also presents data on the half-life and tenth-life of deviations from long-run means. In addition to the country specific results, the last row includes cross-section averages. On the whole, the individual country tests do not provide support for PPP, a result that is consistent with the extant literature. For the model including constant terms, the unit root null can be rejected in only 14 of 51 tests, and it can be rejected in 10 out of 51 tests that use the constant plus trend deterministic specification.

It can also be seen that the average half-life of PPP deviations is 4 years, and the average tenth-life is 12 years. The range of half-lives of PPP deviations is quite wide, from 1 to 17 years, and the range of tenth-lives is even wider, from 2 to 64 years. While the average half-life of deviations from PPP for advanced countries is only 3½ years, it is 5⅓ years for developing countries. Tenth-lives of PPP deviations are 50 percent higher in developing countries, relative to advanced countries. Moreover, 18 out of the 24 rejections of the unit root null are in advanced countries. A possible explanation is that on the whole, advanced countries tend to be more open to foreign trade, an environment which is more conducive to rapid arbitrage in tradables markets. This suggests that one is more likely to reject the unit root null, and to find support for PPP, in a sub-sample of advanced countries. Thus, the panel tests will be applied to the full sample of countries as well as sub-samples of advanced and developing countries.

For six countries, the tests are able to reject nonstationarity under both specifications, namely Algeria, Chile, Israel, the Netherlands, Norway, and Singapore. This group of countries includes high inflation economies, such as Israel and Chile, and highly open economies, such as Singapore and the Netherlands.[12] These results suggest that high inflation and openness are characteristics associated with a more rapid erosion of deviations from PPP. First, unless there are severe trade and exchange restrictions, the exchange rate in high inflation economies will need to adjust frequently to prevent a severe loss of competitiveness. Indeed, several medium to high inflation economies have adopted crawling pegs. Second, arbitrage in tradables markets clearly has a more significant macroeconomic role in relatively

---

[10]Appendix II provides details on the construction of the indices.

[11]For a classification of countries, see the IMF's World Economic Outlook, October 1998, Statistical Appendix, Table B.

[12]The average annual inflation rate in 1971-1997 was 48 percent in Chile and 49 percent in Israel. In the same period the average ratio of foreign trade to GDP was 83 percent in the Netherlands and 265 percent in Singapore.

Table 1. Real Effective Exchange Rates: Unit Root Tests and Related Indicators

Country	Constant			Constant and Trend			Correction of deviations from long-run average 2/	Correction of deviations from long-run average 2/
	ADF/DF	T	Lag length	ADF/DF	T	Lag length	50 percent	90 percent
Algeria	-2.80*	21	5	-3.67**	21	5	3	5
Argentina	-2.58	26	0	-2.53	26	0	2	4
Australia	-0.76	22	4	-4.04**	25	1	11	43
Austria	-2.54	26	0	-2.95	24	2	3	11
Belgium	-3.10**	25	1	-3.14	25	1	2	3
Bolivia	-1.51	26	0	-2.26	26	0	4	8
Brazil	-1.90	26	0	-1.89	26	0	3	13
Canada	-1.94	25	1	-2.95	25	1	3	12
Chile	-3.94***	25	1	-3.93**	25	1	6	20
Colombia	-1.46	25	1	-2.02	23	3	4	7
Denmark	-2.92*	25	1	-3.23*	25	1	2	3
Ecuador	-0.79	26	0	-2.56	24	2	4	6
Egypt	-2.19	26	0	-2.19	26	0	1	2
Finland	-1.89	24	2	-1.36	24	2	2	3
France	-2.78*	25	1	-3.10	25	1	2	3
Germany	-3.05**	25	1	-3.05	25	1	2	2
Greece	-2.94**	23	3	-2.31	23	3	5	9
India	-0.25	26	0	-1.95	24	2	9	26
Indonesia	-0.36	26	0	-1.94	26	0	11	36
Iran	-0.90	26	0	-1.73	26	0	4	12
Ireland	-1.84	22	4	-3.08	22	4	3	6
Israel	-4.12***	23	3	-4.09**	23	3	2	3
Italy	-1.39	26	0	-1.84	26	0	3	5
Japan	-1.31	22	4	-3.22	25	1	17	64
Korea	-2.99**	26	0	-2.92	26	0	2	3
Malaysia	-0.28	26	0	-2.66	25	1	17	58
Mexico	-2.57	26	0	-2.75	26	0	1	2
Morocco	-1.27	25	1	-0.86	25	1	6	11
Netherlands	-3.25**	25	1	-3.84**	22	4	1	2
New Zealand	-3.33**	25	1	-3.18	25	1	2	3
Nigeria	-2.20	25	1	-2.44	25	1	2	3
Norway	-3.69**	25	1	-3.43*	25	1	2	3
Pakistan	-2.55	26	0	-3.38*	26	0	11	35
Peru 1/	-0.31	26	0	-1.56	26	0	na.	na.
Phillipines	-1.71	26	0	-1.70	26	0	3	8
Portugal	-2.57	25	1	-2.26	25	1	2	4
Saudi Arabia	-1.60	25	1	-3.67**	25	1	6	18
Singapore	-3.00**	25	1	-3.35*	25	1	5	18
South Africa	-1.52	26	0	-2.42	26	0	3	11
Spain	-2.37	26	0	-1.72	26	0	3	10
Sri Lanka	-2.43	26	0	-1.04	26	0	2	4
Sweden	-1.78	26	0	-2.38	22	4	2	3
Switzerland	-1.97	26	0	-1.90	26	0	1	2
Taiwan	-1.92	26	0	-2.75	26	0	7	23
Thailand	-1.21	26	0	-2.75	25	1	16	55
Tunisia	-1.61	26	0	0.00	26	0	5	12
Turkey	-1.47	26	0	-2.06	26	0	4	11
United Kingdom	-1.95	25	1	-3.03	25	1	2	5
United States	-2.73*	25	1	-2.68	25	1	2	2
Uruguay	-2.27	25	1	-2.30	25	1	2	2
Venezuela	-1.69	23	3	-2.52	23	3	4	7
Average	-2.07	25	1	-2.56	25	1	4	12

Source: Authors' calculations.

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%;

1/ Lag polynomial with unstable roots;
2/ In years. Based on 5th order autoregression.

Table 2. IPS Tests for the Real Exchange Rate

(Model includes constant term)

	IPS Psi adjusted statistic	Average t statistic from ADF regressions	N
Unadjusted data			
Full sample	-2.29**	-1.73	51
Advanced economies 1/	-3.20	-2.05	25
Developing economies	-0.10	-1.42	26
Low inflation economies	-1.86*	-1.76	28
High inflation economies	-1.36	-1.68	23
Open economies	-2.51**	-1.91	25
Closed economies	-0.75	-1.55	26
Group-demeaned data			
Full sample	-1.02	-1.54	51
Advanced economies 1/	-4.77***	-2.36	25
Developing economies	-1.17	-1.64	26
Low inflation economies	0.21	-1.36	28
High inflation economies	-2.79***	-1.98	23
Open economies	-2.09**	-1.82	25
Closed economies	-0.58	-1.52	26

Source: Authors' calculations.

Note: *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

1/ For definition of country groups see text and Table A1.

Table 3. IPS Tests for the Real Exchange Rate

(Model includes constant and trend terms)

	IPS Psi adjusted statistic	Average t statistic from ADF regressions	N
Unadjusted data			
Full sample	-2.73**	-2.37	51
Advanced economies 1/	-2.61***	-2.51	25
Developing economies	-1.26	-2.24	26
Low inflation economies	-2.79***	-2.51	28
High inflation economies	-0.97	-2.20	23
Open economies	-2.52**	-2.49	25
Closed economies	-1.35	-2.26	26
Group-demeaned data			
Full sample	3.13***	-2.43	51
Advanced economies 1/	-2.51**	-2.49	25
Developing economies	-1.12	-2.23	26
Low inflation economies	-2.59***	-2.48	28
High inflation economies	-1.97**	-2.41	23
Open economies	-1.81*	-2.35	25
Closed economies	-1.38	-2.28	26

Source: Authors' calculations.

Note: *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

1/ For definition of country groups see text and Table A1.

open economies. The panel tests were therefore also be applied to sub-samples of high inflation and low inflation economies, and closed and open economies.[13]

Clearly, there is some overlap between these country groups, as most advanced economies are open and have had low inflation rates, while developing countries tend to be relatively closed and have had higher inflation. Nevertheless, the overlap is not perfect; and it seems interesting to investigate whether quantifiable country characteristics can impinge on the results of tests for PPP.

Table 2 presents IPS t-bar tests for the model that includes a constant but no trend term. The results in the upper part of Table 2 are found using unadjusted data series, while those in the lower part use group de-meaned data. The null hypothesis is that all real exchange rate series under consideration contain a unit root, and the alternative hypothesis is that some of these real exchange rate series are stationary.

The test results for the full sample highlight the importance of de-meaning to remove cross-sectional correlation, which will typically be attributable to a large degree to worldwide economic shocks. Whereas tests based on the original data indicate a rejection of the null, the test based on de-meaned data indicates that the nonstationarity null cannot be rejected. Moreover, the test results seem to depend significantly on country characteristics. Specifically, results for de-meaned data show that the unit root null can be rejected for advanced countries, for open economies, and for high inflation economies, although it cannot be rejected for the full sample.

Table 3 presents the same panel tests for the model that includes a constant term and a time trend. In these tests, which are conventional in the literature, the unit root null is rejected for the full sample with both the unadjusted and the de-meaned data. The results for de-meaned data confirm that real exchange rates behave differently in advanced and open economies, as opposed to developing and closed economies. The distinction between low and high inflation economies, on the other hand, does not appear significant. It should be noted, however, that including a trend constitutes a weakening of the traditional Casselian notion of PPP. The results in Table 2, and particularly those using the demeaned data, may more adequately reflect the usual notion of PPP.

---

[13]Countries in which inflation averaged more than 10 percent annually during 1971-1997 were classified as high inflation. This cutoff is approximately equal to the median inflation rate. There is also some evidence that inflation begins to negatively affect growth at about this rate (Sarel, 1996). Openness was measured in the first instance by the share in GDP of exports plus imports, averaged over the period 1971-1997. The median openness, on this measure, was 38 percent for developing countries and 51 percent for advanced countries. In order to control for the size of the economies, a cross-section regression of the openness indices on GDP levels was run. Countries for which the residuals of this regression were positive (negative) were classified as relatively open (closed).

In sum, the panel tests provide significant support for the PPP hypothesis, in particular for sub-samples of advanced and open economies. It seems, however, that PPP is not a particularly relevant assumption for developing or relatively closed economies. Even so, care must be taken not to read too much into these statistical findings, which should be taken as indicative but not definitive. Even if the real exchange rate is found to have a unit root, this may simply reflect too short a sample period to observe long-run reversion; and, conversely, when a unit root is rejected, this may show only that shocks to the exchange rate were relatively quiescent over the sample period.

III. Long-run Exchange Rate Models

As seen in the previous section, the evidence in favor of the purchasing power parity hypothesis is mixed. Some countries and country groups seem to satisfy PPP, to a reasonable approximation, while others do not. If real exchange rates are indeed non-stationary and there are systematic long-run deviations from PPP, it is natural to ask whether there are stable long-run relationships between some other variables and the real exchange rate.

There is by now an extensive literature on this question; Rogoff (1996) provides a recent survey.[14] Three explanations for long-run deviations from PPP are most often cited: the Balassa-Samuelson hypothesis, differences in the size of the public sector, and changes in the terms of trade.

The Balassa-Samuelson hypothesis is probably the best-known and most influential of these three explanations. It holds that price levels are higher in advanced countries, and their real exchange rate is consequently more appreciated, because the tradables sector is relatively more productive and the relative price of tradables is lower. Following the exposition in Obstfeld and Rogoff (1996), let the domestic and foreign price indices be geometrically weighted averages of traded and nontraded goods and services, where the price of tradables is taken as numeraire and set to 1:

$$P = (1)^{\gamma}\, p^{1-\gamma} = p^{1-\gamma}, \qquad P^{*} = (1)^{\gamma}\, (p^{*})^{1-\gamma} = (p^{*})^{1-\gamma} \tag{2}$$

It follows that the ratio of the price levels in the two countries depends only on the ratio of nontradables prices:

$$\frac{P}{P^{*}} = \left(\frac{p}{p^{*}}\right)^{1-\gamma} \tag{3}$$

[14]Contributions since this survey was written include Chinn (1997) and Ito, Isard, and Symansky (1997).

Taking logarithms and differentiating, and assuming that the (linear homogenous) production functions for tradables and nontradables are given by $y_T = A_T F(K_T, L_T)$ and $y_N = A_N G(K_N, L_N)$ in both the domestic and foreign economies, one obtains

$$\hat{P} - \hat{P}^* = (1-\gamma)(\hat{p} - \hat{p}^*) = (1-\gamma)\left[\frac{\mu_{LN}}{\mu_{LT}}(\hat{A}_T - \hat{A}_T^*) - (\hat{A}_N - \hat{A}_N^*)\right] \tag{4}$$

where $\mu_{LN} / \mu_{LT} \geq 1$ is the ratio of the shares of labor income in the non-tradables and tradables sectors. This expression implies that to the extent that countries have grown rich primarily through improvements in productivity in the tradeables sector, the Balassa-Samuelson hypothesis may be tested by relating a country's real exchange rate to the relative growth of its aggregate productivity, or per capita real GDP.

Differences in government spending provide a second explanation for long-run deviations from PPP. The standard argument here is that higher public spending shifts demand toward the non-tradeables sector, which would raise the domestic price level and lead to an appreciation of the real exchange rate. This effect will of course be transitory unless higher public spending gives rise to some other distortions with long-run real effects, for example because such spending is financed by distortionary taxes. Indeed, there is considerable evidence that a larger public sector and a higher tax burden are associated with decreased aggregate productivity growth (Barro and Lee, 1994). Of course, if higher government spending leads to much lower long-run growth of per capita GDP, the effect on the real exchange rate could also be negative. Another route by which higher public spending may affect the real exchange rate is via larger fiscal and external current account deficits.

Finally, an improvement in the terms of trade may bring about an appreciation of the real exchange rate. The wealth effect associated with such an improvement will tend to lead to an increase in consumption, which—at least in a small country for which tradeables prices are given by the world market—will raise the price of non-tradeables and hence the domestic price level.

Statistical methods

The focus in this study is on the long-run behavior of exchange rates, and the statistical methods used will therefore need to be geared toward allowing for the estimation of consistent long-run parameters in a regression of the exchange rate on the potential explanatory variables: real per capita GDP, the terms of trade, and government consumption relative to GDP.

It has become conventional to view long-run parameters as reflecting cointegrating relationships among a set of I(1) variables. The standard estimation methodology first establishes the order of integration of the variables in question, using one or more of a large number of tests. Having established that the variables may be characterized as I(1), further tests are used to ascertain whether there is at least one linear relationships among these variables such that the residual emerging from the long-run regression is I(0). The most commonly used test of order of integration is based on the Augmented Dickey-Fuller (ADF) statistic; cointegration among variables found to be I(1) is then established using the Engle-Granger two-step procedure, the Johansen procedure, or other such method.

This paper takes a different approach, for essentially two reasons. First, there are as yet few, (and even fewer statistically satisfactory) tests of cointegration in a panel data context.[15] Moreover, it is well known by now that tests of the order of integration do not reliably distinguish between series that contain a unit root and those that are stationary with a "near-unit" root. Second, it has been recently shown by Pesaran and Shin (1998) that long-run parameters may be consistently estimated using the traditional autoregressive-distributed lag (ARDL) approach, provided that sufficiently high lag orders are chosen. Indeed, it turns out that the ARDL approach yields consistent and asymptotically normal estimates of the long-run coefficients irrespective of whether the underlying regressors are I(1) or I(0); and that it compares favorably in Monte Carlo experiments with conventional methods of cointegration analysis.

Against this background, the parameter estimates presented in this section were obtained using two recently developed methods for the statistical analysis of dynamic panel data: mean group (MG) and pooled mean group (PMG) estimation. These methods are particularly suited to the analysis of panels which have both large time and cross-section dimensions ("data fields").

MG estimation derives the long-run parameters for the panel from an average of the long-run parameters from ARDL models for individual countries (Pesaran and Smith, 1995). For example, if

$$\alpha_i(L)\, y_{it} = \beta_i(L)\, x_{it} + \gamma_i\, z_{it} + \epsilon_{it} \tag{5}$$

is the ARDL model for country i, where $i=1,\dots N$, then the long-run parameter for country i is

$$\phi_i = \frac{\beta_i(1)}{\alpha_i(1)} \tag{6}$$

and the MG estimate for the panel as a whole is

[15]A panel test using a null of cointegration has been developed by McCoskey and Kao (1998). There are also a number of (as yet unpublished) residual-based tests of cointegration in a panel setting, for example Kao (1997) and Pedroni (1995).

$$\phi = \frac{1}{N} \sum_{i=1}^{N} \hat{\phi}_i \tag{7}$$

It can be shown that MG estimation with sufficiently high lag orders yields consistent (in fact, superconsistent) estimators of the long-run parameters even when the regressors are I(1).

PMG estimation occupies an intermediate position between MG estimation, in which both the slopes and the intercepts are allowed to differ by country, and classical fixed effects estimation, in which the slopes are fixed and the intercepts are allowed to vary. In PMG estimation, only the long-run coefficients are constrained to be the same across countries, while the short-run coefficients are allowed to vary (Pesaran, Shin, and Smith, 1998). Suppose one wishes to estimate an ARDL model for $i=1,...,N$ countries and $t=1,...,T$ time periods

$$y_{it} = \sum_{j=1}^{p} \lambda_{ij} y_{i,\, t-j} + \sum_{j=0}^{q} \delta'_{ij} \boldsymbol{x}_{i,\, t-j} + \mu_i + \epsilon_{it} \tag{8}$$

where $\boldsymbol{x}_{ij}$ is the vector of regressors for group i, μ_i represents the fixed effects, and the λ_{ij} and $\boldsymbol{\delta}_{ij}$ are parameters (bold notation indicates vectors or matrices). T must be large enough to allow separate estimation for each group. This model can be reparametrized and the time-series observations for each group stacked into vectors

$$\Delta \boldsymbol{y}_i = \phi_i \boldsymbol{y}_{i,\,-1} + \boldsymbol{X}_i \boldsymbol{\beta}_i + \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta \boldsymbol{y}_{i,\,-j} + \sum_{j=0}^{q-1} \Delta \boldsymbol{X}_{i,\,-j} \boldsymbol{\delta}^*_{ij} + \mu_i \boldsymbol{\iota} + \boldsymbol{\epsilon}_i \tag{9}$$

Assuming that the model is stable and hence $\phi_i < 0$, there exists a long-run relationship between y_{it} and x_{it} which is given by

$$y_{it} = -\frac{\boldsymbol{\beta}'_i \boldsymbol{x}_{it}}{\phi_i} + \eta_{it} = \boldsymbol{\theta}'_i \boldsymbol{x}_{it} + \eta_{it} \tag{10}$$

for each $i=1,...N$, where η_{it} is a stationary process. The assumption of long-run homogeneity is then simply

$$\boldsymbol{\theta}_i = \boldsymbol{\theta}, \quad i = 1,...,N \tag{11}$$

Estimation of this model is by maximum likelihood. It may be shown that under some regularity assumptions, the parameter estimates in this model are consistent and asymptotically normal for both stationary and I(1) regressors.

Both MG and PMG estimation require selecting appropriate lag orders for the individual country equations. This selection was made using the Akaike information criterion: the procedure considers all possible lag orders on all variables (up to a maximum lag order) and selects the specification with the best AIC value.

In addition, the results of MG and PMG estimation can be biased by the presence of common factors. It is therefore conventional to also provide estimates on data expressed as deviations from the group mean in each time period when the existence of such common factors is suspected. The most important common factor during the floating exchange rate period, which is covered by this study, was the variation in oil prices, which increased sharply in connection with the oil shocks of 1973 and 1979, and which declined markedly in the mid-1980s.

Although MG and PMG estimation will provide consistent parameter estimates, the interpretation of the parameters will depend on the true order of integration of the underlying process, as pointed out by Fisher and Seater (1993). As the order of integration of the variables considered in this paper is ambiguous on empirical grounds, one will need to look to theoretical arguments. For example, as discussed earlier, one may have a priori reasons to believe that the *level* of the real exchange rate would be related to the *level* of the terms of trade, even if it is difficult to conclusively settle that the former is I(0).

Estimation results

Three models are estimated. The first focuses exclusively on the Balassa-Samuelson effect by modeling the real effective exchange rate as a function of relative real GDP per capita (details of the calculation of this index and the others used in this study are provided in Appendix II). The second model includes the terms of trade and relative government consumption in addition to relative real GDP per capita. In the third model, the dependent variable is the nominal effective exchange rate, and the relative price index is included among the explanatory variables. This specification is analogous to the one employed in Canzoneri, Cumby and Diba (1996). It allows for a separate measurement of the effect of relative prices on nominal exchange rate movements. A unit coefficient on the relative price term would indicate that in the long run, relative price movements are passed through completely to the nominal exchange rate.

All of the models are estimated using the full sample of countries, and for subsamples comprising the advanced and developing countries. In addition, it was assumed that all of the long-run coefficients are the same, and not just a subset. By and large, this hypothesis is not rejected by the data.

Table 4. Pooled Mean Group and Mean Group Estimates - Model 1

(Dependent variable: LRER)

	Pooled-Mean-Group Estimates			*Mean-Group Estimates*			*Hausman Test*	
	Coef.	St. Er.	t-ratio	Coef.	St. Er.	t-ratio	h	p-val
				Unadjusted Data				
All countries								
LYCR	0.312	0.045	6.920	-3.126	3.698	-0.845	0.860	0.350
Advanced countries								
LYCR	0.248	0.050	5.004	-6.715	7.532	-0.892	0.850	0.360
Developing countries								
LYCR	0.965	0.075	12.895	0.326	0.538	0.606	1.440	0.230
				Group-Demeaned Data				
All countries								
LYCR	0.555	0.056	9.867	-0.217	1.149	-0.189	0.450	0.500
Advanced countries								
LYCR	0.251	0.055	4.550	5.227	4.360	1.199	1.300	0.250
Developing countries								
LYCR	0.119	0.066	1.801	2.802	2.519	1.112	1.130	0.290

Source: Authors' calculations.

Note: RER denotes the real effective exchange rate, NER the nominal effective exchange rate, YCR effective relative per capita real income, TOT the terms of trade, GCR the ratio of government consumption to GDP, and LPR the relative price index The operator L denotes the natural logarithm.

Table 5. Pooled Mean Group and Mean Group Estimates - Model 2

(Dependent variable: LRER)

	Pooled-Mean-Group Estimates			Mean-Group Estimates			Hausman Test	
	Coef.	St. Er.	t-ratio	Coef.	St. Er.	t-ratio	h	p-val
				Unadjusted Data				
All countries								
LYCR	0.358	0.033	10.927	0.454	0.504	0.900	0.040	0.850
LTOT	0.322	0.045	7.217	0.028	0.463	0.060	0.410	0.520
GCR	-0.006	0.002	-2.456	-0.041	0.037	-1.097	0.870	0.350
Advanced countries								
LYCR	0.319	0.034	9.276	0.298	0.575	0.519	0.000	0.970
LTOT	0.316	0.059	5.327	-0.138	0.592	-0.233	0.590	0.440
GCR	-0.004	0.003	-1.459	-0.032	0.046	-0.698	0.370	0.540
Developing countries								
LYCR	0.620	0.125	4.965	0.604	0.832	0.727	0.000	0.980
LTOT	0.232	0.066	3.514	0.187	0.717	0.261	0.000	0.950
GCR	0.046	0.005	8.607	-0.049	0.059	-0.835	2.610	0.110
				Group-Demeaned Data				
All countries								
LYCR	0.487	0.050	9.826	-0.403	0.802	-0.502	1.240	0.270
LTOT	0.277	0.046	6.033	0.261	0.267	0.976	0.000	0.950
GCR	-0.010	0.004	-2.619	-0.082	0.085	-0.971	0.730	0.390
Advanced countries								
LYCR	0.504	0.050	10.155	0.601	0.670	0.896	0.020	0.880
LTOT	0.937	0.070	13.304	0.400	0.478	0.837	1.290	0.260
GCR	-0.004	0.003	-1.186	-0.056	0.044	-1.254	1.370	0.240
Developing countries								
LYCR	0.127	0.067	1.902	-0.234	0.417	-0.562	0.770	0.380
LTOT	0.077	0.032	2.421	0.201	0.243	0.824	0.260	0.610
GCR	0.059	0.003	20.751	-0.044	0.049	-0.909	4.470	0.030

Source: Authors' calculations.

Note: Variable definitions are as in Table 4.

Table 6. Pooled Mean Group and Mean Group Estimates - Model 3

(Dependent variable: LNER)

	Pooled-Mean-Group Estimates			*Mean-Group Estimates*			*Hausman Test*	
	Coef.	St. Er.	t-ratio	Coef.	St. Er.	t-ratio	h	p-val
				Unadjusted Data				
All countries								
LPR	0.981	0.003	339.700	1.100	0.382	2.882	0.100	0.760
LYCR	0.390	0.050	7.866	-2.031	2.090	-0.971	1.340	0.250
LTOT	0.355	0.035	10.053	0.237	0.418	0.567	0.080	0.780
GCR	-0.002	0.002	-1.011	0.028	0.071	0.394	0.180	0.670
Advanced countries								
LPR	0.738	0.031	23.683	0.875	0.353	2.477	0.150	0.700
LYCR	-0.072	0.061	-1.189	-3.439	2.884	-1.192	1.360	0.240
LTOT	0.266	0.062	4.292	-0.607	0.754	-0.805	1.350	0.250
GCR	-0.041	0.005	-7.677	-0.031	0.081	-0.377	0.020	0.900
Developing countries								
LPR	0.982	0.003	296.092	1.316	0.673	1.957	0.250	0.620
LYCR	0.611	0.088	6.926	-0.677	3.053	-0.222	0.180	0.670
LTOT	0.263	0.047	5.616	1.048	0.324	3.238	6.020	0.010
GCR	0.003	0.004	0.903	0.084	0.115	0.733	0.500	0.480
				Group-Demeaned Data				
All countries								
LPR	0.971	0.003	301.405	1.434	0.626	2.293	0.550	0.460
LYCR	0.259	0.035	7.475	-1.162	1.776	-0.655	0.640	0.420
LTOT	0.304	0.031	9.923	1.218	0.484	2.516	3.580	0.060
GCR	-0.003	0.003	-1.060	-0.200	0.177	-1.130	1.240	0.260
Advanced countries								
LPR	0.969	0.005	176.386	1.047	0.311	3.370	0.060	0.800
LYCR	0.534	0.044	12.258	-1.484	1.378	-1.076	2.150	0.140
LTOT	1.268	0.074	17.037	1.283	1.436	0.894	0.000	0.990
GCR	0.015	0.003	4.673	-0.071	0.054	-1.311	2.540	0.110
Developing countries								
LPR	0.982	0.003	321.546	0.775	0.094	8.280	4.940	0.030
LYCR	0.254	0.052	4.841	0.698	0.622	1.123	0.520	0.470
LTOT	0.160	0.033	4.788	0.075	0.156	0.480	0.310	0.580
GCR	0.027	0.004	6.935	0.049	0.048	1.019	0.210	0.650

Source: Authors' calculations.

Note: Variable definitions are as in Table 4.

Table 4 summarizes the results for Model 1. The estimates provide support for the Balassa-Samuelson hypothesis. The PMG coefficients are positive and significant for both the advanced and developing countries, and for the overall sample. The estimated elasticity becomes smaller for the developing countries when the group-demeaned data are used. As indicated earlier, this primarily reflects the presence of a common factor. Oil prices are an obvious candidate—developing countries were significantly more affected than advanced countries by the oil shocks of the 1970s owing inter alia to the much higher share of oil imports in their GDP (Gelb, 1986). The MG estimates provide less information than the PMG estimates, with coefficients that are of varying sign and typically insignificant. However, the Hausman tests reported in the last two columns of the table indicate that the data do not reject the restriction of common long-run coefficients, so pooling the data (by using the PMG estimator) would appear to be an acceptable and informative procedure.

The estimation results for Model 2 are shown in Table 5. The PMG estimates once again provide ample support for the Balassa-Samuelson hypothesis, for both advanced and developing countries, and the pooling restrictions cannot be rejected. There is also considerable evidence of a positive terms of trade effect on the real exchange rate. As would be expected, this effect is considerably attenuated for the developing countries when the demeaned data are used. The evidence of an effect of government spending on the real exchange rate is weaker and somewhat mixed. The estimated coefficients tend to be negative and (borderline) insignificant for the advanced countries. For the developing countries, by contrast, the PMG coefficients are positive and significant, but the Hausman tests reject the pooling restrictions for the demeaned data.

As indicated earlier, Model 3 treats separately the nominal and relative price components of the real effective exchange rate. As Table 6 shows, usually about 75 to 100 percent of relative price movements are passed though to the nominal effective exchange rate in the long run. Support for the Balassa-Samuelson effect remains strong, except, and surprisingly, for the advanced countries when unadjusted data are used. The other results are little changed compared with Model 2. It is worth noting in this connection that although the pooling restriction is rejected for the terms of trade effect in the developing countries (using unadjusted data), both the MG and the PMG estimates are positive and significant.

For the sake of completeness, panel unit root tests were conducted on the explanatory variables used in the regressions. The relative real GDP per capita indices appear to be I(1), as do the nominal effective exchange rate and (with one exception) the relative price indices. By contrast, the terms of trade are clearly I(0), while the relative government consumption ratios appear to by I(0) for the advanced and low-inflation countries.

IV. CONCLUSION

The principal conclusion of the paper is that PPP holds to a considerable extent, but that other factors, notably relative growth performance and the terms of trade, also affect the real exchange rate in the long run. While the direction of causality is of course subject to debate, the central point is that these variables do in most cases exhibit a marked positive long-run correlation with the real exchange rate, as would be predicted on the basis of the theoretical arguments advanced in this paper. This contrasts somewhat with earlier studies of the real exchange rate, which have not always found such strong support for the Balassa-Samuelson hypothesis and terms of trade effects.

These results would appear to be applicable to the estimation of the appropriate level of the real exchange rate following a large macroeconomic shock. For example, it may be argued that the sharp exchange rate depreciations experienced by several Asian countries during the crisis of 1997-98 in part reflected a downward revision in investors' assessment of their growth performance. Indeed, real GDP contracted sharply in the affected economies. However, the elasticities estimated in this paper provide support for the view that the very large real depreciations initially observed in these economies were excessive. Even with an elasticity on relative real GDP per capita of ½, a permanent downward revision in potential GDP of, say 20 percent, would have been associated with a steady-state decline in the real exchange rate of just 10 percent. Not surprisingly, once the initial shock passed, real exchange rates in the affected economies recovered to a considerable extent (nominal exchange rates of course, showed a less pronounced recovery, as the exchange rate depreciation also fueled higher inflation).

References

Balassa, B. (1964), "The Purchasing Power Parity Doctrine: A Reappraisal," *The Journal of Political Economy* 72, 584-596.

Barro, R.J. and J. Lee, 1994, "Sources of Economic Growth," *Carnegie-Rochester Conference Series on Public Policy*.

Breuer, J. B., 1994, "An Assessment of the Evidence on Purchasing Power Parity," in: *Estimating Equilibrium Exchange Rates*, John Williamson (ed.), Institute for International Economics.

Canzoneri, M. B., R. E. Cumby, and B. Diba, 1996, "Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries," *NBER Working Paper Series* No. 5676, National Bureau of Economic Research.

Chinn, M. D., 1997, "Sectoral Productivity, Government Spending and Real Exchange Rates: Empirical Evidence for OECD Countries," *NBER Working Paper Series* No. 6017, National Bureau of Economic Research.

Coakley, J. and A. M. Fuertes, 1997, "New Panel Unit Root Tests of PPP," *Economics Letters* 57, 17-22.

Dornbusch, R., 1988, *Exchange Rates and Inflation*, MIT Press.

Edwards, S. and M. A. Savastano, 1998, "Exchange Rates in Emerging Economies: What do We Know? What do We Need to Know?" unpublished manuscript, Research Department, International Monetary Fund.

Fisher, M.E. and J.J. Seater, 1993, "Long-Run Neutrality and Superneutrality in an ARIMA Framework," *American Economic Review* 83, 402-415.

Frankel, J., 1986, "International Capital Mobility and Crowding Out in the U.S. Economy: Imperfect Integration of Financial Markets or Goods Markets," in: *How open is the US economy*, R. Hafer (ed.), Lexington Books.

Frenkel, J., 1981, "The Collapse of Purchasing Power Parity During the 1970s ," *European Economic Review* 16, 145-165.

Froot, K. A. and K. Rogoff, 1994, "Perspectives on PPP and Long-Run Real Exchange Rates," *NBER Working Paper Series* No. 4952, National Bureau of Economic Research.

Gelb, A.H., 1986, "Adjustment to Windfall Gains: A Comparative Analysis of Oil-Exporting Countries," in J. P. Neary and S. van Wijnbergen (eds.), *Natural Resources and the Macroeconomy*, MIT Press.

Heston, A., and R. Summers, 1991, "Penn World Table (Mark 5): An Expanded Set of International Comparisons 1950-1988," *Quarterly Journal of Economics* 106, 327-368.

Im, K. S., M. H. Pesaran, and Y. Shin, 1997, "Testing for Unit Roots in Heterogeneous Panels," revised version of *DAE Working Papers Amalgamated Series* No.9526, University of Cambridge.

Ito, T., P. Isard, and S. Symansky, 1997, "Economic Growth and Real Exchange Rate: An Overview of the Balassa-Samuelson Hypothesis in Asia," *NBER Working Paper Series* No. 5979, National Bureau of Economic Research.

Krugman, P., 1978, "Purchasing Power Parity and Exchange Rates," *Journal of International Economics* 8, 397-407.

MacDonald, R., 1996, "Panel Unit Root Tests and Real Exchange Rates," *Economics Letters* 50, 7-11.

McCoskey, S. and C. Kao, 1998, "A Residual-Based Test of the Null of Cointegration in Panel Data," *Econometric Reviews*, 17(1), 57-84.

Mesquita, M., 1996, *Determinants of the Equilibrium Exchange Rate in the Brazilian Economy*, D.Phil. dissertation, Oxford University.

Nagayasu, J., 1998, "Does the Long-Run PPP Hypothesis Hold for Africa?: Evidence from Panel Co-Integration Study," *Working Paper* 98/123, International Monetary Fund, Washington, D.C.

O' Connell, P.G.J., 1998, "The Overvaluation of Purchasing Power Parity," *Journal of International Economics 44*, 1-19.

Obstfeld, M. and K. Rogoff, 1996, *Foundations of Open-Economy Macroeconomics*, MIT Press.

Pesaran, M.H. and Y. Shin, 1998, "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis," in S. Strom, A. Holly, and P. Diamond (eds.), *Centennial Volume of Ragnar Frisch*, Cambridge University Press.

Pesaran, M. H., Y. Shin, and R. P. Smith, 1998, "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels," unpublished manuscript, University of Cambridge.

Pesaran, M. H. and R. Smith, 1995, "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics* 68, 79-113.

Rogoff, K., 1996, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, XXXIV, 647-668.

Sarel, M., 1996, "Nonlinear Effects of Inflation on Economic Growth," *Staff Papers* 43, International Monetary Fund, 199-215.