Chapter 32

PERSPECTIVES ON PPP AND LONG-RUN REAL EXCHANGE RATES

KENNETH A. FROOT

Harvard and NBER

and

KENNETH ROGOFF*

Princeton University

Contents

*The authors are grateful to Marianne Baxter, Jose De Gregorio, Rudiger Dornbusch, Hali Edison, Jeffrey Frankel, Gene Grossman, Karen Lewis, Richard Marston, John Rogers and Julio Rotemberg for helpful comments, and to Michael Kim and Paul O'Connell for research assistance.

Handbook of International Economics, vol. III, Edited by G. Grossman and K. Rogoff

1. Introduction

This paper overviews what we know – and what we don't know – about the long-run determinants of purchasing power parity. A decade ago, when the papers for the first edition of the *Handbook of International Economics* were written, PPP seemed like a fairly dull research topic. On the one hand, the advent of floating exchange rates made it obvious to even the most stubborn defenders of purchasing power parity that PPP is not a short-run relationship; price level movements do not begin to offset exchange rate swings on a monthly or even annual basis. On the other hand, there were neither sufficient time spans of floating rate data nor adequate econometric techniques for testing the validity of PPP as a long-run relationship. Fortunately, the past decade has witnessed a tremendous degree of progress in the area and, in spite of some mis-steps and research tangents, several important results have emerged.

First, a broad body of evidence suggests that the real exchange rate is not a random walk, and that shocks to the real exchange rate damp out over time, albeit very slowly. Consensus estimates put the half-life of deviations from PPP at about 4 years for exchange rates among major industrialized countries. Second, there is some evidence that real exchange rates tend to be higher in rich countries than in poor countries, and that relatively fast-growing countries experience real-exchange rate appreciations. But the empirical evidence in favor of a "Balassa–Samuelson" effect is weaker than commonly believed, especially when comparing real exchange rates across industrialized countries over the post-Bretton Woods period.

Section 2 of the paper reviews the huge time series literature testing simple PPP. This area has proven fruitful ground for applying modern methods for dealing with nonstationary and near-nonstationary time series. Our organization traces out the evolution of the literature, from naive static tests of PPP, to modern unit-root approaches for testing whether real exchange rates are stationary, to cointegration techniques, the most recent phase of PPP testing. As we shall see, cointegration approaches have sometimes created as much confusion as clarity on the issue of PPP. It appears that this approach is plagued by small-sample bias when applied to floating exchange rates, often yielding nonsensical results.

Because convergence to PPP is relatively slow, it is not easy to empirically distinguish between a random-walk real exchange rate and a stationary real exchange rate that reverts very slowly. This is particularly problematic when looking at highly volatile floating exchange rates, where the noise can easily mask slow convergence toward long-run equilibrium. One of the major innovations has been to look at longer historical data sets, incorporating fixed as well as floating rate periods. There are some obvious problems in mixing regimes, though these have been addressed to some extent recently. One issue that has not been looked at in the literature is the problem of selection bias, which we discuss in Section 2.3.6.

In Section 2.5 we discuss research on more disaggregated price data, including a nearly two-hundred year data set on commodity prices in England and France during the seventeenth and eighteenth centuries. Aside from providing an extremely long data set, this historical data offers some perspective on the behavior of cross-country relative prices in more modern times.

In Section 3 of the paper, we look at some possible medium- to long-run determinants of the real exchange rate, particularly the supply-side determinants emphasized in the popular Balassa–Samuelson model. We also consider some evidence that positive demand shocks, such as unexpected increases in government spending, lead to medium-run appreciations of the real exchange rate. Finally, we consider "pricing-to-market" theories. The conclusions offer some possible directions for future research.

2. Evolving tests of simple PPP

In this section, we examine simple PPP – Cassel's (1922) notion that exchange rates should tend to equalize relative price levels in different countries.[1] While this notion appears simple enough, many subtleties arise in trying to implement it. In Subsection 2.1, we begin by briefly reviewing the basic motivation underlying PPP and some alternative definitions. Following that, we turn in the next three subsections to the very large recent literature on testing for PPP. We distinguish three different stages of PPP tests:

(1) older tests in which the null hypothesis is that PPP holds (Section 2.2)

(2) more recent theories and time series tests in which the null hypothesis is that PPP deviations are completely permanent (Section 2.3); and

(3) even more recent cointegration tests in which the null hypothesis is that deviations away from *any* linear combination of prices and exchanges rates is permanent (Section 2.4).

We show how each stage reflects reactions to prior empirical results as well as to advances in theoretical modeling and econometric technique. Finally, in Section 2.5, we consider tests based on more disaggregated price data.

2.1. Definitions and basic concepts

The starting point for most derivations of PPP is *the law of one price*, which states that for any good i,

[1] See Dornbusch (1987) for a historical treatment of the PPP doctrine.

$$p_t(i) = p_t^*(i) + s_t \tag{2.1}$$

where $p_t(i)$ is the log of the time-t domestic-currency price of good i, $p_t^*(i)$ is the analogous foreign-currency price, and s_t is the log of the time-t domestic-currency price of foreign exchange. The premise underlying the law of one price is a simple goods-market arbitrage argument: abstracting from tariffs and transportation costs, unfettered trade in goods should ensure identical prices across countries. In practice, the "law" of one price holds mainly in the breech, as we shall later see. Still, it provides a very useful reference point.

If the law of price holds for every individual good, then it follows immediately that it must hold for any identical basket of goods.[2] Most empirical tests, however, do not attempt to compare identical baskets, but use different countries' CPIs and WPIs instead. In general, these have weights and mixes of goods that vary across countries. [In principle, it is possible to construct international price indices for identical baskets of goods and, starting in the 1950s, there have been a few attempts to do so. These culminate in the influential work of Summers and Heston (1991), discussed later in Section 3.]

Absolute consumption-based PPP requires:

$$p_t(\mathrm{CPI}) = p_t^*(\mathrm{CPI}) + s_t \tag{2.2}$$

where CPI denotes the basket of goods used in forming the consumption price index. Clearly, even if the law of one price holds, there is no reason why condition (2.2) should hold, unless the two countries have identical consumption baskets. In order to allow for a constant price differential between baskets, the bulk of the empirical literature focuses on testing *relative consumption-based PPP*:

$$\Delta p_t(\mathrm{CPI}) = \Delta p_t^*(\mathrm{CPI}) + \Delta s_t \tag{2.3}$$

which requires that *changes* in relative price levels be offset by changes in the exchange rate. Indeed, much of the post World War I debate over re-establishing pre-war parities, which provided the genesis of PPP theory, implicitly referred to relative PPP. Of course, among low inflation economies, there is little more reason to believe that (2.3) will hold than (2.2), since real shocks can lead to changes in the relative prices of different goods baskets. Across countries with very different inflation rates, however, one might expect condition (2.3) to hold even when (2.2) does not.

Indeed, much of economists' faith in PPP derives from a belief that over most of the past century, price level movements have been dominated by monetary factors. If price index movements are dominated by monetary shocks, and if

[2] Even if the law of one price fails for individual goods, it is possible that the deviations roughly cancel out when averaged across a basket of goods.

money is neutral in the long run, then it won't matter if the two baskets being compared are not the same; relative PPP should still hold (approximately). Of course, economists like to use PPP as a frame of reference not just for hyperinflationary economies, but for any pair of economies. Most of this section will be concerned with straightforward tests of PPP, but later in Section 3, we shall consider various adjustments that have been proposed to try to give PPP more meaning for low-inflation economies.

2.2. Stage one: Simple purchasing power parity as the null hypothesis

In Cassel's (1922) view, PPP was seen as a central tendency of the exchange rate, subject to temporary offset, and not a continuously-holding equivalence. Much of the work on PPP through the 1970s [see Officer's classic (1976a) survey] recognizes the importance of temporary disturbances to PPP, in principle. But early formal empirical analyses were limited by the absence of statistical and theoretical tools for distinguishing between short-run and long-run real effects. Thus, typically, the early studies at best only allowed for a disturbance term, and did not specifically allow for any dynamics of adjustment to PPP.

Without doubt, the most positive results in stage-one tests came from data on high inflation economies. Frenkel (1978) ran regressions of the form

$$s_t = \alpha + \beta(p_t - p_t^*) + \varepsilon_t \tag{2.4}$$

for a number of hyperinflationary economies. He was not so much interested in the properties of the error term, as in whether the slope coefficient was one. Frenkel indeed found estimates of β quite close to one and, based on these estimates, argued that PPP should be an important building block of any model of exchange rate determination.

Outside of hyperinflations, however, most stage-one tests produced strong rejections of PPP. (Today, of course, it is well known that stationarity of the residuals in eq. (2.4) is required for standard hypothesis testing, a condition that will fail if some types of shocks to the real exchange rate are permanent.) Frenkel (1981) reports that PPP performed poorly for industrialized countries during the 1970s, with β estimates typically far from one (some country-pairs actually yield negative coefficients while for others β estimates exceeded 2.0). Frenkel suggested that the failure of PPP might be attributable to some combination of temporary real shocks and sticky goods prices, implicitly arguing that PPP still holds in the long run even though short-run factors get in the way of finding $\beta = 1$. However, Frenkel made no attempt to model the short-run bias in the coefficients.

Aside from failing to allow for dynamic adjustment, another obvious problem with eq. (2.4) is that exchange rates and prices are simultaneously determined, and there is no compelling reason to put exchange rates on the left-hand side, rather than visa-versa. Indeed, many authors [e.g. Isard (1977) and Giovannini (1988)] ran the reverse regression, projecting relative prices on the exchange rate.

Krugman (1978) was an attempt to explicitly address the endogeneity problem [see also Frenkel (1981)]. Krugman offered a flex-price model which had the domestic monetary authorities offsetting the effects of real shocks by expanding the money supply and thereby raising the price level. Krugman showed that in this case the endogeneity of the price level introduces a downward bias in OLS estimates of β in eq. (2.4). To control for this bias, Krugman (1978) and Frenkel (1981) re-estimated the equation using instrumental variables.[3] Their methodology succeeded in that it yielded coefficients closer to one than under OLS, though one could still soundly reject purchasing power parity. The endogeneity issue can, of course, also be cast as a left-out regressor problem. That is, the bias in the key coefficient β can be removed by conditioning the regression on the real exogenous factors that affect both exchange rates and prices and which, according to some model, explain deviations from PPP. We will look at some of these factors later in Section 3.

A fundamental flaw in the econometrics of stage-one tests was the failure to take explicitly into account the possible nonstationarity of relative prices and exchange rates. Today it is well known that if there is a unit root in the error term to eq. (2.4), then standard hypothesis tests of the proposition $\beta = 1$ are invalid. Both the stage-two and stage-three tests we consider next are explicitly designed to deal with this problem. Overall, the main lesson from stage-one tests was that PPP does not hold continuously, but the results provided no perspective on whether PPP might be valid as a long-run proposition.

2.3. *Stage two: The real exchange rate as a random walk*

Stage-one tests' disappointing results and flawed hypothesis testing led to an alternative approach. In stage-two tests, the null hypothesis becomes that the real exchange rate follows a random walk, with the alternative hypothesis being that PPP holds *in the long run*. These tests stand those from stage-one tests on their head: they impose – rather than estimate – the hypothesis that $\beta = 1$, and test – rather than impose – the hypothesis that the (log of the) real exchange rate

$$q_t \equiv s_t - p_t + p_t^* \tag{2.5}$$

is stationary. Examples of early stage-two tests include Darby (1983), Adler and Lehman (1983), Hakkio (1984), Frankel (1986), Huizinga (1987) and Meese and

[3] Krugman (1978) used a time trend as an instrument, whereas Frenkel (1981) used a time polynomial as well as lagged exchange rates and price levels.

Rogoff (1988). As we shall see, the main problem with stage-two tests is low power. Given the phenomenal volatility of floating exchange rates, it can be very hard to distinguish between slow mean reversion and a random walk real exchange rate, especially if one relies only on post-Bretton Woods data. Much of the evolution of stage-two testing has revolved around finding longer or broader data sets, and implementing more powerful unit roots tests.

Leaving aside the problem of low power, how plausible is the null that the real exchange rate follows a random walk? Roll (1979) argued that a random walk is a sensible null hypothesis because real exchange-rate changes, like changes in asset prices, should not be predictable if foreign exchange markets are efficient. Of course, this analogy is inappropriate, since real exchange rates are not traded assets and therefore not subject to the usual efficient capital markets logic. Indeed, there is no reason why even the nominal exchange rate – which is a market variable – should follow a random walk in the presence of nominal interest differentials or risk premia.

Certainly it is possible to find rationales for random walk, or near random walk, exchange rate behavior that are more defensible than Roll's. In Section 3 below, we will discuss the Balassa–Samuelson model, in which cross-country sectoral differences in productivity growth can lead to real CPI exchange rate changes. If productivity differential shocks are permanent, sectoral productivity shocks can induce a unit root in the real exchange rate. We also discuss Rogoff's (1992) model, in which intertemporal smoothing of traded goods consumption can lead to smoothing of the *intra*temporal price of traded and nontraded goods. This in turn implies a unit root in the real exchange rate, even when productivity shocks are temporary. Obstfeld and Rogoff (1995) offer a model in which any shock (even a monetary one) that effects a wealth transfer across countries will lead to a potentially long-lasting change in relative work effort, and therefore the real exchange rate.[4] Space considerations prevent us from presenting these and other related rationales for random walk real exchange rates in any detail. For our purposes here, though, it is enough to note that there are a variety of simple yet reasonable models that can generate highly persistent deviations from PPP.

2.3.1. *Econometric techniques to test for random walk real exchange rates*

Once the null hypothesis posits that the real exchange rate follows a random walk (or more generally has a "unit root" component),[5] it becomes necessary

[4]There is a substantial empirical literature on the effects of wealth re-distributions on the long-run equilibrium exchange rate; [see for example, Krugman (1990), and Bayoumi, Clark, Symansky and Taylor (1994)]. For further discussion of the effects of wealth transfers on the long-run equilibrium exchange rate, see Baxter's, and Obstfeld and Rogoff's chapters in this Handbook.

[5]If the real exchange rate has one unit root, then its first difference must be stationary though not necessarily serially uncorrelated as in the random walk model.

to negotiate a number of important econometric subtleties. Most importantly, conventional confidence intervals calculated under the null of a stationary real exchange rate are no longer appropriate and, as Dickey and Fuller (1979) emphasized, the correct confidence intervals should be wider.

The modern literature uses three main techniques for distinguishing the real exchange rate from a random walk.[6] The first, and most commonly used, is the Dickey–Fuller and augmented Dickey–Fuller tests. These involve a regression of the real exchange rate, q_t, on a constant, a time trend, q_{t-1}, and lagged changes in q_{t-1}:

$$q_t = \alpha_0 + \alpha_1 t + \alpha_2 q_{t-1} + \Phi(L)\Delta q_{t-1} + \varepsilon_t \tag{2.6}$$

where L is the lag operator, $\Phi(L)$ is a pth order polynomial in L, with coefficients $\phi_1, \phi_2, \ldots \phi_p$, and ε_t is white noise. Under the null hypothesis that q_t has a unit root, $\alpha_2 = 1$. Under the alternative hypothesis that PPP holds in the long run, $\alpha_1 = 0$ and $\alpha_2 < 1$.[7] The distribution of the OLS estimates for eq. (2.6) is nonstandard under the random walk null, with the appropriate confidence intervals reported by Dickey and Fuller (1979). An example of a study applying the Dickey–Fuller test to floating real exchange rates is Meese–Rogoff (1988), who are unable to reject the unit root hypothesis for monthly dollar/pound, dollar/yen, and dollar/DM floating exchange rate data.

Equation (2.6) can also be used to calculate Phillips (1987) Z test, which allows for conditional heteroskedasticity of the residual. Perron (1989) extends these tests to allow for one-time changes in the constant and the trend by including dummy variables. However, in introducing break points, data snooping biases can make the resulting test statistics difficult to interpret [see Christiano (1992)].

The second commonly-used technique is that of variance ratios. The idea here is that under the null hypothesis of a random walk, the variance of the real exchange rate should grow linearly over time. This implies that the statistic

$$k(i) = \frac{T}{T-i+1} \cdot \operatorname{var}\left[(1-L)\, q_t\right] i/\operatorname{var}\left[\left(1-L^i\right) q_t\right], \; i = 2, 3, \ldots T-1 \tag{2.7}$$

should be one for all i. For a stationary series, on the other hand, the k statistic converges to zero as k increases.[8]

[6]See also Breuer (1994) for an excellent survey of econometric problems in testing for unit roots in real exchange rates. Some of the very early efforts to test the random walk real exchange rate hypothesis, including Darby (1983), and Adler and Lehman (1983), did not use modern root testing methodologies, but nevertheless illustrated the difficulties in rejecting the random walk model.

[7]Some of the studies below test only $\alpha_2 < 1$, and do not jointly apply the restriction $\alpha_0 = 0$. Also, many studies look only at the straight Dickey–Fuller test and do not augment the regression with the lagged changes. There is no problem with this simplification as long as the residuals are not autocorrelated.

[8]Poterba and Summers (1986) show that the variance ratio is a function of the processes' autocorrelation coefficients 1 through i.

A third technique is that of fractional integration, which encompasses a broader class of stationary processes under the alternative hypothesis. A fractionally integrated process allows the real exchange rate to evolve according to:

$$\Phi(L)(1-L)^d q_t = \chi(L)\varepsilon_t, \tag{2.8}$$

where $\Phi(L)$ and $\chi(L)$ are polynomial lag operators with roots outside the unit circle and ε_t is white noise. If the parameter $d = 0$, then the real exchange rate is confined to the class of stationary ARMA processes described by $\Phi(L)$ and $\chi(L)$. If $d = 1$ and $\Phi(L) = \chi(L) = 1$, then the real exchange rate follows a random walk. The advantage of this class of processes is that it allows for fractional integration, $0 < d < 1$. Because fractionally integrated processes are stationary, but have autocovariance functions that die off more slowly than ARMA processes, encompassing them under the alternative hypothesis may enhance one's chances of rejecting the random walk null. [See Diebold, Husted and Rush (1991) for citations and a discussion of estimation techniques.]

2.3.2. *Results for post-Bretton Woods data*

The basic result in the empirical literature is that if one applies unit roots tests to bilateral industrialized-country monthly data, it is difficult to reject the null of a unit root for currencies that float against each other. [See, for example, Meese and Rogoff (1988) or Mark (1990).] An exception is Huizinga (1987), who constructed variance ratios to argue in favor of *positive* autocorrelation in U.S. dollar real exchange rates for horizons under two years. However, Huizinga's results may be attributable to the long, large swings in the dollar between the mid-1970s and mid-1980s.

For currency pairs that are fixed (or formally stabilized), the evidence is more mixed. In Mark's (1990) tests for the 1973–1988 period, the intra-European exchange rates come closest to rejecting a random walk, although it is only for the Belgium/Germany currency pair that a random walk can be rejected at the 5 percent confidence level. Chowdhury and Sdogati (1993) look at the 1979–1990 period, during which time the EMS was in place. They strongly reject the random walk for bilateral rates of various European currencies against the Deutsche mark, but not for European exchange rates against the U.S. dollar. The apparent systematic differences in the behavior of the real exchange rate for various floating versus fixed exchange rates has been noted and explored by a number of authors [see, for example, Mussa (1986) and the Frankel and Rose chapter in this Handbook].

2.3.3. *Power against persistent alternatives*

The major concern with the early stage-two tests of the random walk hypothesis is that they lack sufficient power to reject. Because slow, albeit positive, rates of reversion toward PPP are plausible in many models, random walk tests may provide little information against relevant alternative hypotheses.

To see how important the issue of power is, and to gain a sense of how much data is needed to reject plausible alternatives, it is useful to calibrate a simple autoregression, as done by Frankel (1986, 1990). The results of this analysis show that the post-Bretton Woods sample period is far too short to reliably reject the random walk hypothesis.

Suppose that PPP indeed holds over the long run, and that deviations from PPP follow an AR(1) process (on monthly data), with serial correlation coefficient ρ and error variance σ^2. That is

$$q_t - \overline{q} = \rho(q_{t-1} - \overline{q}) + \varepsilon_t \tag{2.9}$$

where $0 \leq \rho < 1$, $\overline{q}$ is the long-run equilibrium real exchange rate and ε is a white noise error term with variance σ^2. Suppose that the autoregression is run on a panel data set with T observations and N independent bilateral exchange rates, each governed by the same stochastic process.

In these circumstances, the variance of the OLS estimate of ρ is given by[9]

$$\mathrm{var}(\hat{\rho} - \rho) = \sigma^2 / [NT \cdot \mathrm{var}(q - \overline{q})] \tag{2.10}$$

where $\mathrm{var}(q - \overline{q}) = \sigma^2/(1 - \rho^2)$. Thus, we find that the standard error of the OLS estimate converges to[10]

$$p\lim\,[\mathrm{std}\,(\rho - \overline{\rho})] = \left[\frac{1 - \rho^2}{NT}\right]^{1/2} \tag{2.11}$$

How many years of data does expression (2.11) imply one needs to be able to reject the random walk process when the real exchange rate is governed by the stationary process (2.9)? Suppose for a moment that the true half-life of PPP deviations is 36 months (3 years). This translates into a true value of the AR coefficient in monthly data of $\rho = 0.981 = 0.5^{1/36}$.

Assuming 18 years of data ($T = 216$) on a single exchange rate ($N = 1$), eq. (2.11) then implies that the standard error of the OLS estimate of ρ is approximately $0.0132 = \left[(1 - 0.981^2)/216\right]^{1/2}$. With this degree of imprecision, the true value of ρ ($= 0.981$) is only approximately 1.44 [$= (1 - 0.981)/0.0132$)] standard errors away from one. Thus 18 years of data are not likely to be sufficient for rejecting the random walk null – and this calculation uses conventional stationary real exchange rate standard errors, rather than Dickey–Fuller standard errors.

[9]This example assumes for simplicity that the mean of the log real exchange rate, $\overline{q}$, is known. Estimation of the mean can induce finite sample bias in the estimated autoregressive coefficient, but this nuance is not central to our example here.

[10]Note that by using the asymptotic standard error, we avoid small sample problems which introduce non-normality into the distribution of the t-statistic for ρ.

How many years of data would it take to reject $\rho = 1$ at a 5 percent confidence interval for a single currency using the large-sample Dickey–Fuller critical t value of 2.89? Solving the condition

$$2.89^2 = T(1-\rho)^2/(1-\rho^2)$$

implies $T = 864$ months, or 72 years! Obviously, with a longer half life (i.e. a larger value of ρ), even more data would be required. Indeed, the preceding calculation understates the problem, since we have employed asymptotic standard errors in making these calculations.

Two approaches to dealing with the power problem have been tried in the literature; one is to look at a number of currencies simultaneously (allow for $N > 1$), and the other is to look at long-horizon data sets encompassing both pre- and post-Bretton Woods data.[11]

2.3.4. Tests using cross sections of currencies

With 18 years of data, simultaneously testing $N = 4$ *independent*, identically-distributed currencies would expand the data set sufficiently to reject (since $18 \cdot 4 = 72$). Hakkio (1984) was the first to suggest using cross-section data to gain power; he employed GLS to allow for cross-exchange rate correlation in the residuals in four exchange rates against the dollar. Despite the enhanced power of his test, Hakkio was unable to reject the random walk model.

Abuaf and Jorion (1990) perform similar tests, running autoregressions in levels using GLS for ten countries's currencies against the US dollar over the period 1973–1987. The longer time series and the larger cross-section does generate more power, but nevertheless permits only the weakest of rejections of the random walk hypothesis – at the 10 percent significance level using one-sided tests. These results roughly fit our calibration above: with 14.5 years of data and (say) 5 independent bilateral exchange rates, we have the equivalent of 72 years of data. Thus, even with this size cross-section, one would expect rejections to be marginal. It would be interesting to see if adding more recent data to their sample would lead to more decisive rejections.

In an interesting recent study, Cumby (1993) makes clever use of the *Economists'* "Hamburger Standard", which each year reports the dollar price of McDonald's Big Mac hamburgers in up to 25 countries. Although only about 7 years (1987–1993) of data are available, Cumby finds that the large cross-section yields enough power to detect substantial reversion toward the law of one price. In

[11] Some improvement in power can be achieved simply by avoiding inefficient test specifications. Abuaf and Jorion (1990), for example, note that the early tests performed by Adler and Lehman (1983) – which estimated autoregressions of real exchange rate *changes* – were likely to be much less powerful than similar tests performed on the levels of the real exchange rate. Cheung and Lai (1993c) apply a more powerful version of the Dickey–Fuller test due to Elliot, Rothenberg and Stock (1992).

fact, deviations from Big Mac parity exhibit remarkably little persistence, with only 30 percent of the deviation in one year persisting to the next. This fact seems striking given that a large fraction of the 'goods' embodied in a Big Mac, including local infrastructure costs and labor, are essentially nontraded.

How can Cumby's finding of relatively rapid convergence to "hamburger PPP" be reconciled with most other studies of PPP, which find relatively slow rates of convergence? One factor may be that relatively few of the currency pairs in Cumby's sample were actually floating against one another. As we have already seen, convergence appears easier to detect in fixed rate than in floating rate data. Second, peso problems may lead to understated standard errors: the Big Mac sample includes a number of relatively high inflation countries – Argentina, Brazil, Hungary, Malaysia, Mexico, Russia and Thailand – whose currencies are generally pegged except for the occasional large realignment.[12] Finally, McDonald's own pricing policies may produce a more rapid rate of convergence in Big Mac prices than in broader aggregate price indexes.

2.3.5. *Tests using longer time series*

The second approach to improving power is to extend the sample period. Frankel (1986), for example, uses 116 years (1869–1984) of data for the dollar/pound real exchange rate. (Other, earlier, long-horizon studies such as Lee (1976), did not incorporate modern unit root methodology.) Frankel finds that a simple first-order autoregression yields a coefficient of 0.86, which implies that PPP deviations have an annual decay rate of 14 percent and a half life of 4.6 years. His rejection of the unit root null is significant at the 5 percent level, using Dickey–Fuller confidence intervals. Another early attempt to use long samples to test convergence towards PPP is Edison (1987), who looks at dollar/pound data for the years 1890–1978. Edison uses an error-correction mechanism [see also Papell (1994)], regressing the change in the log of the nominal exchange rate, Δs_t, on the contemporaneous change in the log of relative prices, $\Delta(p-p^*)_t$, and the lagged real exchange rate, $(s-p-p)_{t-1}$:

$$\Delta s_t = \alpha_0 + \alpha_1[\Delta(p-p^*)_t] + \alpha_2(s-p-p)_{t-1} \tag{2.12}$$

Edison estimates $\alpha_2 = 0.09$, i.e. that the nominal exchange rate decays towards PPP at a statistically-significant 9 percent per year, implying a half life of roughly 7.3 years. In a similar exercise, Johnson (1990) uses 120 years of Canadian dollar/U.S. dollar exchange rate data. He, too, is able to reject the random walk hypothesis, and finds a half life for PPP deviations of 3.1 years.

Abuaf and Jorion (1990) use time series data from 1901–1972 for eight currencies. Their point estimates suggest a half-life of PPP deviations of 3.3 years, and they are easily able to reject a random walk. Both their results and Frankel's are

[12] See Karen Lewis's chapter in this Handbook for a discussion of the peso problem.

consistent with the simple model calibrated above. Glen (1992) uses variance ratios to test for mean reversion in the real exchange rate for 9 bilateral exchange rates over the 1900–87 period. Glen, too, finds strong evidence of mean reversion.

It must be emphasized that in addition to extending the sample, the long-sample studies discussed above all combine relatively low variance pre-Bretton Woods exchange rate data with the highly volatile post-Bretton Woods data. For the simple first-order AR process specified in (2.9), the variance of the real exchange rate does not affect the power of the test. If, however, the real exchange rate is better described by a richer ARMA process, and if there are different parameters governing fixed versus floating rates, the test results may be heavily affected by the inclusion of fixed rate periods. Thus these papers leave unresolved the question of whether mean reversion would be detected in 100 years of floating rate data. [Wars may also affect PPP dynamics; see Rogers (1994).]

Lothian and Taylor (1994) is an interesting attempt to cast some light on this issue. Their data set consists of almost two centuries of data for the dollar/pound (1791–1990) and franc/pound (1803–1990) real exchange rates. Using only the post-Bretton Woods portion of the data, they are not able to reject the random walk hypothesis for either exchange rate. But when the entire sample is used, the random walk null is easily rejected for either rate. (They estimate half lives of 4.7 years for the dollar/pound and 2.7 years for the franc/pound.) Moreover, using a simple Chow test to compare first-order AR coefficients before and after Bretton Woods, they find that one cannot reject the hypothesis of no structural change. In fact, if one estimates a simple AR (1) model on the pre-Bretton Woods data, it outperforms a random walk model on post-Bretton Woods data at one- to five-year horizons. Thus Lothian and Taylor conclude that there is no evidence for the view that the inclusion of fixed-rate periods biases unit roots tests of the real exchange rate.

Of course, there is at least one striking difference between fixed and floating regimes: Under fixed rates, deviations from PPP must be eliminated by domestic price level movements. The error correction specification (2.12) is ideally suited to measure the degree to which reversion toward PPP occurs through the nominal exchange rate versus through prices. Under floating rates, both Edison (1987) and Johnson (1990) cannot reject the hypothesis that all of the reversion towards PPP is due to exchange rate movements.

Finally, we note two recent studies that have allowed for the possibility that long-run real exchange rate data is better characterized by fractionally integrated processes rather than by the usual ARMA models. Diebold, Husted and Rush (1991) look at a novel data set that encompasses over a hundred years of gold-standard data, and find that little power is added by allowing for fractional integration. They are able to strongly reject the random walk model, but in most cases their estimates suggest that a simple ARMA model best describes real exchange rates.

Cheung and Lai (1993a) arrive at a somewhat different conclusion using a similar technique. Using a shorter time sample (1914–72) than DHR, they are unable to reject a random walk model against fractional and ARMA alternatives. However, they estimate the parameter d in eq. (2.8) to be about 0.5, suggesting evidence of fractional integration in real exchange rates. They also show that the power of fractionally-integrated alternatives to reject a random walk when the true process has $d = 0.5$ is considerably greater than that of standard ARMA alternatives.

2.3.6. *A caveat: Sample selection or "survivorship" bias in long-horizon tests of PPP*

One interesting question that has not previously been raised in the long-sample PPP literature is whether "survivorship" bias might exaggerate the extent to which PPP holds in the long run. Specifically, the countries for which very long-run PPP series are easily available tend to be those few who have continuously been among the world's wealthiest nations. Countries that grew very fast from a low level (e.g. Japan), and countries that were once rich but are no longer so (e.g. Argentina) have not been studied as extensively. But these are precisely the countries for which one might expect the relative price of nontraded goods to have changed most dramatically (see our discussion of the Balassa–Samuelson effect in Section 3), and for which tests of long-run PPP are most likely to fail.

To intuitively gauge the importance of this sample-selection effect, we consider data for the Argentine peso against the US dollar and the British pound over the period 1913–1988. The Argentine CPI and nominal exchange rate data

Table 2.1
Augmented Dickey–Fuller Regressions on Argentine/American and Argentine/British CPI Real Exchange Rates

$$q \equiv s - p + p^*$$
$$q_t = \alpha_o + \alpha_1 q_{t-1} + \phi_1 (1 - L) q_{t-1} + \varepsilon_t$$

Sample Period	Peso/dollar 1913–1988	Peso/pound 1913–1988
N	74	74
α_0	−0.931	−1.466
t-stat (against $\alpha_0 = 0$)	−3.21	−3.14
α_1	0.808	0.764
t-stat (against $\alpha_1 = 1$)	−3.20	−3.12
1% critical value	−3.52	−3.52
10% critical value	−2.59	−2.59
ϕ_1	0.256	0.06
R^2	0.152	0.122
DW	1.92	1.98
σ_q	0.203	0.220

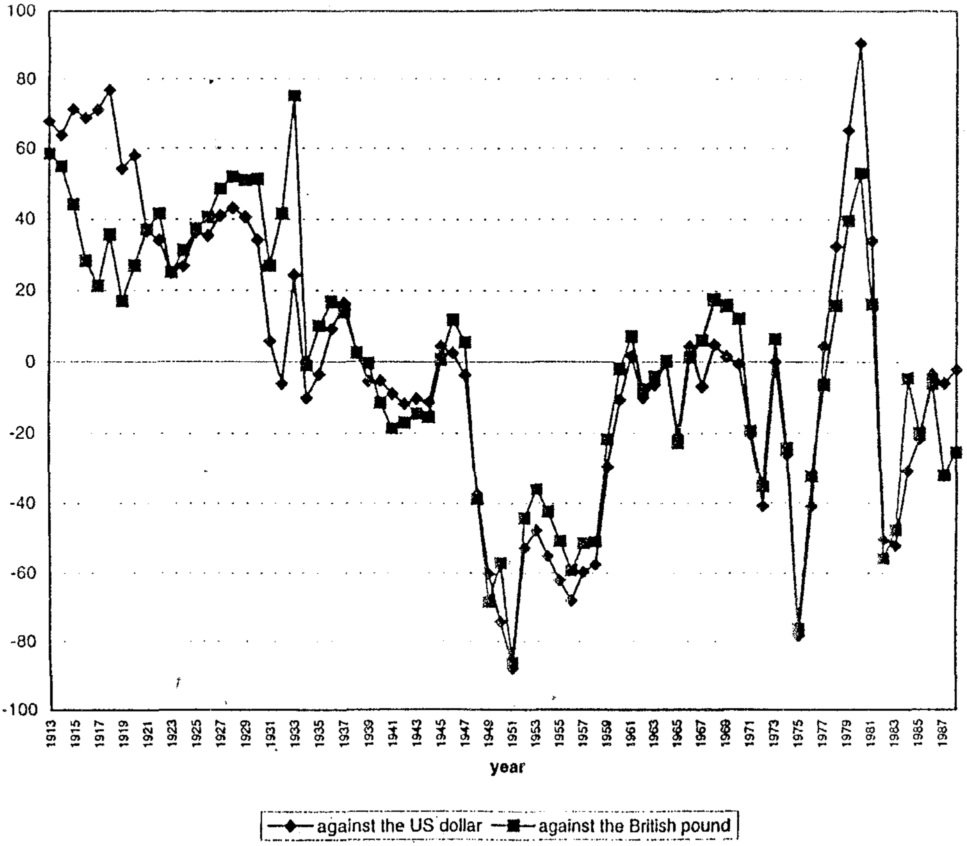

Figure 2.1. Real value of the Argentine peso, 1913–1988 (log percentage deviations from period average).

come from Cavallo (1986), except for the post-1980 data which is from *International Financial Statistics*.

As Figure 2.1 shows, with the exception of the well-known massive overvaluation of the peso during the early 1980s, there is a steady decline of the peso over the period. The real peso has fallen by roughly 80 percent (in log terms) since the beginning of the century, a rate of decline of almost 1 percent per year. This trend is highly statistically significant, as Table 2.1 illustrates.[13] Moreover, the strong rejections of a unit root that emerge in the pound/dollar data are

[13] As the table also illustrates, however, the time trend is not significant under the alternative hypothesis of a stationary real exchange rate.

absent here. Even with seventy-five years of data, it is not possible to reject the hypothesis that the detrended peso/dollar or peso/pound exchange rates follow a random walk.

Although only suggestive, these results indicate that one must be cautious in interpreting results from long-run PPP tests, as the tendency towards long-run PPP may not apply to countries whose incomes relative to the rest of the world have undergone sharp changes.

2.4. *Stage three tests: Cointegration*

At first glance, PPP testing would seem to provide the perfect context for Engle and Granger's (1987) work on cointegration.[14] The techniques are designed to test for long-run equilibrium relationships, for which the adjustment mechanism remains unspecified. Cointegration tests are thus liberated from stage-one concerns about endogeneity and left-out variables. Moreover – and, as we shall see, more controversially – cointegration tests hold forth the promise of testing weaker versions of PPP, since they require only that *some* linear combination of exchange rates and prices be stationary. In other words, stage-two tests ask whether the real exchange rate $q_t \equiv s_t - p_t - p_t^*$ is stationary. Stage-three tests ask only whether

$$s_t - \mu p_t + \mu^* p_t^* \tag{2.13}$$

is stationary for *any* constant μ and μ^*. Any incremental power from stage-three tests over stage-two tests must therefore come from relaxing the symmetry and proportionality restrictions that $\mu = \mu^* = 1$. In the discussion below, we will distinguish between *trivariate* tests that place no restrictions on the coefficients in (2.13), and *bivariate* tests that impose the symmetry restriction $\mu = -\mu^*$.[15]

Why might μ not equal one? Consider the following model used by Taylor (1988), Fisher and Park (1991), and Cheung and Lai (1993a,b). First, assume that PPP holds exactly for traded goods so that

$$s_t = p_t^{\mathrm{T}} - p_t^{*\mathrm{T}} \tag{2.14}$$

where p_t^{T} is the time t (log) home price of traded goods. Second, assume that the overall price index consists of a weighted average of traded and nontraded goods prices:

$$p_t = \gamma p_t^{\mathrm{T}} + (1 - \gamma) p_t^{\mathrm{N}} \tag{2.15}$$

where p_t^{N} is the time t home price of nontraded goods, for which PPP does not

[14] For a review of the cointegraton literature and its applications to macroeconomcs, see Campbell and Perron (1991).

[15] A stage-two test, which imposes $\mu = -\mu^* = 1$, may simply be thought of as the univariate case in this categorization.

necessarily obtain. The price index abroad is similar to (2.15), with weights γ^* and $1-\gamma^*$. Finally, the price of nontraded goods is assumed to be proportional (in the limit) to the price of traded goods:

$$p_t^{\mathrm{N}} = \alpha_0 + \phi p_t^{\mathrm{T}} + \varepsilon_t \tag{2.16}$$

$$p_t^{*\mathrm{N}} = \alpha_0 + \phi p_t^{*\mathrm{T}} + \varepsilon_t^* \tag{2.17}$$

where the residuals ε and ε^* are stationary. Given eqs. (2.14) – (2.17), a regression of the form

$$s_t = \mu p_t + \mu^* p_t^* + \varepsilon_t' \tag{2.18}$$

yields coefficients[16]

$$\mu = \frac{1}{\gamma + \phi(1-\gamma)} \tag{2.19}$$

$$\mu^* = \frac{1}{\gamma + \phi^*(1-\gamma^*)} \tag{2.20}$$

One possible explanation of why the slope coefficients in eqs. (2.16) and (2.17) might not equal one is simply that there is a trend in the relative prices of traded and nontraded goods. Another explanation, offered by Taylor (1988), Fisher and Park (1991), and Cheung and Lai (1993a,b) is that errors in measuring nontraded goods prices can imply ϕ, $\phi^* \neq 1$. But can measurement error interfere with the proportionality of p_t^{N} and p_t^{T}? One possibility is to think of p^{N} as an index of nontraded goods prices that is subject to either "fixed-weight" or "new goods" bias. (For simplicity assume that no such bias exists in the traded-goods price index.) Fixed-weight bias results when fixed-weight price indices confront changing relative prices. Bryant and Cecchetti (1993) show how these effects can generate permanent upward index movements when relative prices change, and therefore bias measured inflation upward. A second source of bias comes from the introduction of new goods, which one can think of having high implicit prices prior to their introduction.

Thus, in principle, it is possible to think of plausible reasons why one might want to allow for $\mu \neq \mu^* \neq 1$ in eq. (2.13). We turn next to giving a brief overview of cointegration methods.

2.4.1. Techniques and potential applications to real exchange rates

Cointegration techniques ask whether a group of nonstationary variables can be combined to produce a stationary variable. If so, the nonstationary variables are said to be cointegrated. More precisely, consider the $N \times K$ matrix X_t,

[16] We are implicitly assuming that the home and foreign price indices are not themselves cointegrated. If they are, one can impose the assumption $\mu = \mu^*$ in the cointegrating regression.

which consists of all the dependent and independent variables in the system. Suppose, for example, that individually, the variables are integrated of order one (i.e. are stationary in first differences, as is the case with exchange rates and prices).[17] Then if there exists a linear combination of the data, given by the $1 \times N$ vector $B(i)$, such that $B(i)X_t$ is stationary, then we say that X_t is cointegrated. Denoting the matrix of all vectors that yield stationary results by B, the rank of B $(r < N)$ gives the number of cointegrating vectors.

Early applications of cointegration methods to testing PPP were based on a three-step procedure. In the first stage, one tests the exchange rate and the two domestic price series for unit roots, using the augmented Dickey–Fuller test as in eq. (2.6) above. For the bivariate case, of course, there are only two series, the exchange rate and relative prices.

Assuming that one cannot reject the random walk hypothesis for any of the variables, the second stage is to estimate the cointegrating regression (2.18) using OLS. For the bivariate case one imposes $\mu = -\mu^*$. (If one can reject the unit root hypothesis for at least one variable, but cannot for at least one other variable, one cannot reject the no-cointegration null.)

Cointegration of prices and exchange rates implies that the error term in eq. (2.18), ε_t, is stationary. Thus, the third step is to use the OLS residuals from (2.18) to run the Dickey–Fuller regression (2.6), but with the time trend omitted, and to test the hypothesis that $\alpha_2 = 1$. Using this approach, prices and exchange rates are *not* cointegrated under the null hypothesis, whereas they are cointegrated under the alternative hypothesis $\alpha_2 < 1$.[18]

The three-step method is inherently inefficient in part because it requires choosing, somewhat arbitrarily, a single right-hand side variable. More recent PPP tests have been able to avoid this inefficiency, using a technique due to Johansen (1991). Johansen proposed a one-step full-information maximum-likelihood estimator for estimating the coefficients in specifications such as eq. (2.18), and simultaneously testing for the presence of a unit root. Unlike the method above, the ML estimates are not influenced by which variable is on the left-hand side of the single equation regression. The parameter estimates are thus more efficient, and the Johansen test for cointegration thus more powerful than a two-step test.[19] Horvath and Watson (1993) extend the

[17] Ogaki and Park (1990) distinguish between "deterministic" and "stochastic" cointegration. The former is satisfied if a linear combination of X_t is stationary around a deterministic trend, whereas the latter requires that the linear combination of X_t contain no trend. Our definition is essentially one of deterministic cointegration.

[18] Fisher and Park (1991) employ a test proposed by Park that takes cointegration to be the null hypothesis and no cointegration to be the alternative. In essence, this test is constructed by adding a time polynomial to the right-hand side of eq. (2.6), and testing its significance.

[19] Cheung and Lai (1993c) provide evidence for the Johansen test's higher power.

Johansen methodology to allow for constraints that represent long-run equilibrium conditions; this effectively transforms the Johansen test into a stage-two procedure.[20]

2.4.2. *Empirical results of cointegrating tests of PPP*

A plethora of studies have applied cointegration methods to testing PPP. A partial list includes Edison and Klovland (1987), Corbae and Ouliaris (1988), Enders (1988), Kim (1990), Mark (1990), Fisher and Park (1991), Cheung and Lai (1993a), and Kugler and Lenz (1993). Surveys of this material include Giovannetti (1992) and Breuer (1994).

These studies reveal several systematic features of the data. First, it is easier to reject the no-cointegration null across pairs of currencies that are fixed than across pairs that are floating. This finding is consistent with the stage-two results discussed in section (2.3). Second, one finds that tests based on CPI price levels tend to reject less frequently than tests based on WPIs. One explanation for this finding is that consumer price indices have a higher nontraded goods component than wholesale prices, which tend to weight manufactured goods more heavily.[21]

A third common finding is that for post-Bretton Woods floating exchange rates, rejections of the no-cointegration null occur more frequently for trivariate systems (where p and p^* enter separately) than for bivariate systems (where they enter as $p - p^*$), or for stage-two tests (where the coefficient on $p - p^*$ is constrained to be one).[22] Weakening the proportionality and symmetry restrictions therefore makes the residuals appear more stationary.

At first glance, these results seem to provide a strong endorsement for stage-three tests (cointegration) over stage-two tests, since they are generally more successful in rejecting the random walk hypothesis. The problem, unfortunately, is that the estimates of μ and μ^* vary wildly across the various studies based on modern floating rate data, and the magnitudes are often rather implausible. Cheung and Lai, for example, find coefficients that range from 1.03 to 25.4 for CPIs, and 0.3 to 11.4 for WPIs, with most of the coefficients coming in above 1. Imposing the symmetry restriction (looking at the bivariate case instead of the trivariate case) reduces this range only slightly.

Rationalizing these extreme empirical estimates of μ is difficult, to say the

[20]For an application of the Horvath and Watson procedure, see Edison, Gagnon, and Melick (1994).

[21]Keynes (1932) sharply criticized Churchill's Exchequer for using WPIs when making PPP calculations to evaluate Britain's decision to return to its pre-World War I gold parity. Keynes argued WPIs were misleading as index of the real exchange rate because they did not sufficiently reflect nontraded goods prices. McKinnon (1971) also argues that PPP should hold to a much greater extent for WPIs than for CPIs.

[22]See, for example, Cheung and Lai (1993a), Mark (1990), and Kugler and Lenz (1993).

least. How large a bias, for example, can be rationalized by the model embodied in eqs. (2.14) – (2.17)? Bryant and Cecchetti (1993) attempt to measure the size of the "weighting bias" by comparing CPI-index inflation with the rate of inflation that emerges as a common component across goods included in the CPI index. They estimate that weighting bias leads to an overstatement of inflation of about 0.6 percent per annum for the CPI and about 0.35 percent per annum for the personal consumption expenditure deflator. Lebow, Roberts and Stockton (1992) attempt to estimate the size of the new goods bias, and find that it leads to an overstatement of inflation by at most 0.5 percent per annum. Thus, taken together, these two sources of measurement error bias might raise consumer price inflation by roughly 1 percent per year. Then, if inflation averages 5 percent, these effects might raise ϕ from 1 to 1.2, implying that $\mu = 0.83$. This of course, assumes that there is no similar bias in the traded goods index, which would push μ back towards one.

Nor can a trend rise in nontraded relative to traded goods prices explain values of μ far from one. Assuming that both monetary factors and productivity differentials are trend stationary, the coefficient in eq. (2.16) turns out to be $\phi = (\lambda_m + \lambda_a)/\lambda_m$, where λ_m and λ_a are the rates of money growth, and traded relative to nontraded goods productivity growth respectively. Thus if we take inflation to 5 percent and the trend traded/nontraded goods productivity growth differential to be 2 percent, then $\phi = 1.4$, and $\mu = 0.71$.

Clearly, it is very difficult to interpret the results of cointegration tests when estimates of the cointegrating vector have no apparent economic meaning.[23] One possible explanation for the wide-ranging coefficient estimates is small-sample bias. Banerjee et al. (1986) show that in finite samples, cointegrating regressions can result in substantial bias, and that the severity of this bias is related to R^2 – they suggest that regressions with $R^2 < 0.95$ are likely to lead to substantial bias. The problem of low R^2 is, of course, especially likely to plague exchange rate regressions over floating rate data.

Indeed, cointegration tests seem to yield much more reliable results when estimated over long sample periods, rather than just over post-Bretton Woods data. Kim (1990), for example, uses WPI and CPI real exchange rates for the US against five countries – Canada, France, Italy, Japan and the United Kingdom – during the 1900–1987 period. He is always able to reject no cointegration, and he finds coefficients that are strikingly close to one in all cases but for that of Canada.[24]

[23] See also Hakkio and Rush (1991) and Breuer (1994) for critiques of unit root and cointegration tests of PPP.

[24] Kim runs the cointegrating regression (2.13) for the CPI and WPI respectively, and finds ϕ coefficients of 0.99 and 0.98 (France), 0.99 and 0.98 (Italy), 1.00 and 0.98 (Japan), 0.96 and 1.00 (United Kingdom), and 0.73 and 0.55 (Canada). The R^2 are high in all these regressions (averaging around 0.96) except for Canada, in accordance with the theoretical results of Banerjee et al. (1986).

2.4.3. *Summary: What have we learned from stage-three tests?*

There have been a plethora of papers applying cointegration testing to PPP, but on the whole it is not clear that technique has yet provided a net benefit over earlier stage-two tests; indeed, it may have produced some misleading results due to small sample bias. Over longer time periods, and for fixed rates, the bias becomes less serious. Thus far, however, the results from cointegration tests on long-horizon data have not produced any insights not available from stage-two tests.[25]

2.5. *Tests using disaggregated price data*

In order to gain a deeper understanding of why PPP fails, a number of studies have attempted to look at a central building block of PPP, the law of one price.

In his classic (1977) paper, Isard looks for, and finds, deviations from PPP where one would least expect them – in highly disaggregated traded goods price indices. He reports large and persistent deviations from the law of one price in U.S. and German export transactions prices for various 2 though 5 digit SITC categories (e.g. pumps, internal combustion engines, etc.) and in U.S. export unit values in 7-digit A and B groupings when compared to similar unit values from Canada, Germany and Japan. Isard goes on to demonstrate a positive correlation between contemporaneous dollar exchange rates and relative dollar prices.[26] He speculates that this correlation might disappear at longer horizons, but (as with other stage-one tests) does not formalize or test this conjecture explicitly.

Giovannini's (1988) paper is similar in spirit. He finds deviations from PPP not only among disaggregated traded goods, but even among basic "commodity" manufactured goods, such as ball bearings, screws, nuts and bolts. Giovannini's data (which come from the Bank of Japan) compare Japanese domestic and export prices (on shipments bound for the US) during the floating rate period. In line with Isard's results, Giovannini finds large and persistent deviations from the law of one price that are strongly correlated with the nominal exchange rate.[27]

[25] Johansen and Julius (1992) argue that deviations from PPP and deviations from uncovered interest parity may be cointegrated, so that it is important to analyze both simultaneously. This presumes, of course, that deviations from PPP have a unit root component.

[26] Of course, such correlations can be trusted only to the extent that the exchange rate and relative prices are stationary.

[27] Isard and Giovannini both suggest that sticky nominal prices may account for the exchange rate/relative price correlation. See Frankel and Rose's paper in this Handbook for a detailed discussion of the effects of nominal exchange rate movements on the real economy.

2.5.1. Disaggregated price data for the modern floating rate period

Several recent studies that employ disaggregated data have investigated the extent to which departures from PPP are caused mainly by the presence of nontraded goods versus deviations from the law of one price in traded goods. To see this dichotimization, suppose the real exchange rate is $q_t \equiv s_t - p_t + p_t^*$ as defined in eq. (2.5), and the price index in each country is a weighted average of traded and nontraded goods prices $p_t = \gamma p_t^{\mathrm{T}} + (1 - \gamma)p_t^{\mathrm{N}}$ as in eq. (2.15). Combining these two expressions, we can write

$$q_t = (s_t - p_t^{\mathrm{T}} + p_t^{*\mathrm{T}}) - (\gamma - 1)(p_t^{\mathrm{T}} - p_t^{\mathrm{N}}) - (\gamma^* - 1)(p_t^{*\mathrm{T}} - p_t^{*\mathrm{N}}) \tag{2.21}$$

so that real exchange rate depends on deviations from the law of one price in traded goods, as well as on the relative price of traded and nontraded goods within each country.

One study that addresses this issue is Engel (1993). Engel examines a multi-country data set of individual prices, including goods of varying degrees of tradedness. Engel finds that monthly fluctuations from the law of one price for individual traded goods across countries are very large in comparison with fluctuations in relative prices within a country. Even for apparently homogenous traded goods such as bananas, the deviations from the law of one price can be large and volatile. Rogers and Jenkins (1995) extend this result in two ways. First, they sort out traded and nontraded components of the CPI and find that, on average, 81 percent of the variance in the real CPI exchange rate is explained by changes in the relative price of traded goods (which they measure using food prices).

Both of these studies seem to support the view that deviations from the law of one price in traded goods – the first term on the right-hand side of eq. (2.21) – dominate short-term real exchange rate fluctuations. One important qualification to these results is that they are based on retail (CPI) data, and even the "traded" goods embody substantial nontraded inputs. The retail price of bananas includes not only the traded goods input, but local shipping, rent and overhead for the retailer, and labor. Indeed, for many seemingly highly-traded goods, these indirect costs can far outweigh direct traded-goods costs.

Engel and Rogers (1994) provide some further perspective on this issue. Their analysis is based on CPI data for both U.S. and Canadian cities for 14 categories of consumer prices. They find that the variability in the price of a good in two different locations within a country depends on the distance (and the squared distance) between locations, as in gravity models of trade. However, they find that holding other variables (including distance) constant, the variability in prices between two U.S. or two Canadian cities is much less than between a Canadian and a U.S. city. Crossing the U.S.–Canadian border adds as much to the variability of prices as adding (a minimum) of 2500 miles between cities within a country.

Engel and Rogers interpret their finding as strong evidence that prices are sticky in local currency, and that changes in the exchange rate lead to deviations in the law of one price. While their evidence is striking, retail goods generally contain substantial nontraded components, and these components may be much larger across countries than within countries. For example, labor may be much more mobile between New York and Los Angeles than between cross-border neighbors such as Buffalo and Toronto.

Rogers and Jenkins (1993) look at the persistence of deviations from PPP for each component of the CPI across 11 OECD countries as well as across 54 disaggregated goods between the US and Canada. For each good (or index component) i, they test whether deviations in the law of one price

$$q_t(i) = s_t - p_t(i) + p_t^*(i) \tag{2.22}$$

follow a random walk, using the augmented Dickey–Fuller test in eq. (2.6) without a time trend. (Thus, this is a stage-two-type test.) Interestingly, for highly nontraded index components (such as rent), Rogers and Jenkins are unable to reject a random walk for any of the 39 country pairs. They occasionally reject, however, when food prices are used as the index. When looking at more disaggregated individual goods prices between the U.S. and Canada over samples which run from the mid-1970s to 1990, they find similar results:

(1) it is not possible to reject a random walk in the relative price of haircuts (a nontraded good); and

(2) rejection rates are considerably higher for potatoes, eggs, etc. (which are taken to be traded goods).

While disaggregation appears quite informative, the papers by Engel and by Rogers and Jenkins all use relatively short sample periods. They may have enough data to detect statistical differences in relative-price variances, but not enough to provide much power to detect differences in persistence.

2.5.2. *Tests using disaggregated price data and longer times series samples*

In order to obtain a longer time series of disaggregated price data and to gain some perspective on the behavior of prices during recent periods, we consider data from the period 1630–1789 in England and France for three commodities: wheat, charcoal and butter.[28] All prices are in terms of silver (implicitly we assume that the law of one price holds for silver.)

[28]The main data sources are Beveridge (1939), Hauser (1936), and Jastram (1981). For further discussion of the data, see Froot, Kim and Rogoff (1995), who look at deviations from the law of one price for a seven-hundred year data set for England and Holland.

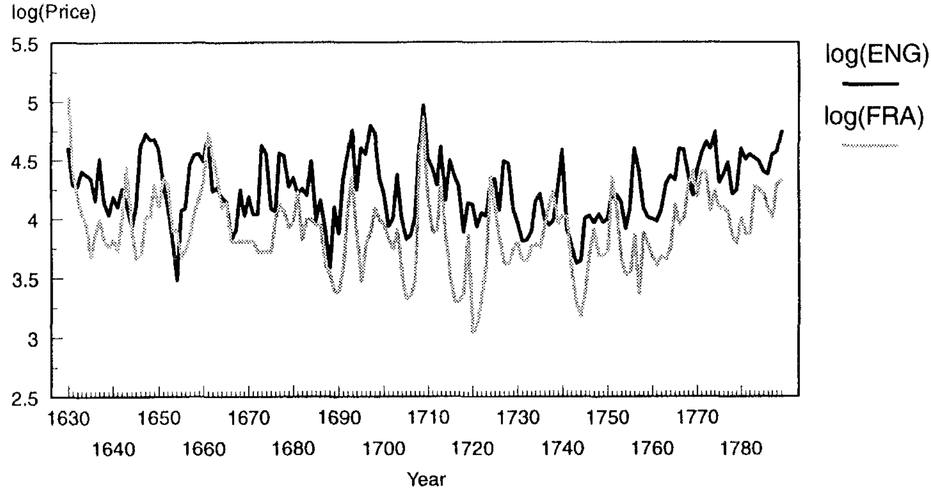

Figure 2.2. Log wheat price, 1630–1789. Price in grams of silver per hectoliter of wheat.

The individual and relative (log) prices are graphed in Figures 2.2–2.4, which reveal several striking aspects of the data. First, note the high volatility of goods prices within a country and of relative prices for the same good across countries. Deviations in PPP for wheat are the most volatile, with a standard deviation of 30 percent per annum. Relative butter and charcoal prices follow, with annual standard deviations of 17 and 12 percent respectively.

The second striking fact is the appearance of trends in individual commodity prices. The log price of wheat, measured in hectoliters per gram of silver and depicted in Figure 2.2, shows little or no trend between 1630 and 1789, averaging about 4.08 during the sample. Indeed, as Froot, Kim and Rogoff (1995) note, this level is strikingly near today's relative price of 4.00.[29] Over a time span this long – almost four centuries – it might be fair to assume that the absence of a relative price trend suggests roughly equal growth rates in wheat and silver productivity.

For charcoal prices, shown in Figure 2.3, there is a slight upward trend over the sample. In addition, charcoal is on average 25 percent more expensive in France than in England over the sample. The price differential probably reflects endowments (England had a greater domestic supply), though the price differential is, of course, bounded by customs charges and transportation costs.

Figure 2.4 depicts the (log) price of butter, measured in kilograms of butter per gram of silver, for the years 1717 to 1789. Butter is probably the least traded of the three goods in our sample, since there was no refrigeration during this period. Note that in contrast to charcoal, butter's price was initially 2.7

[29] This is based on a price \$3.40 per American bushel of wheat and \$5.45 per troy ounce of silver.

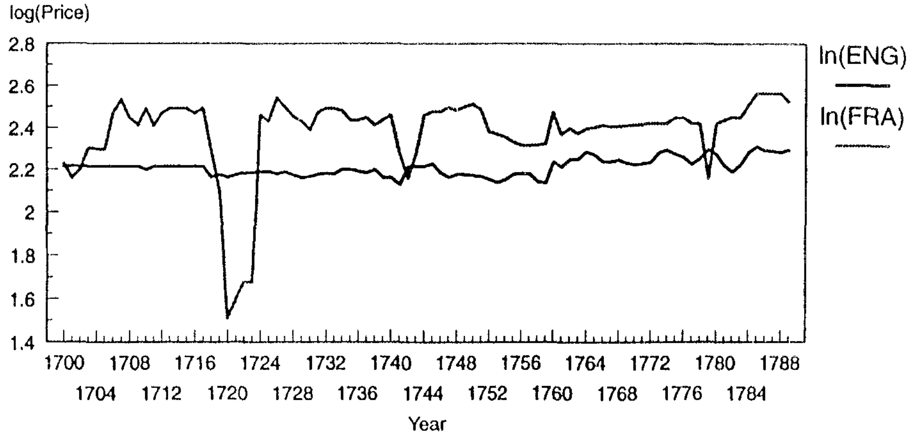

Figure 2.3. Log charcoal prices, 1700–1789. Price in grams of silver per hectoliter of charcoal.

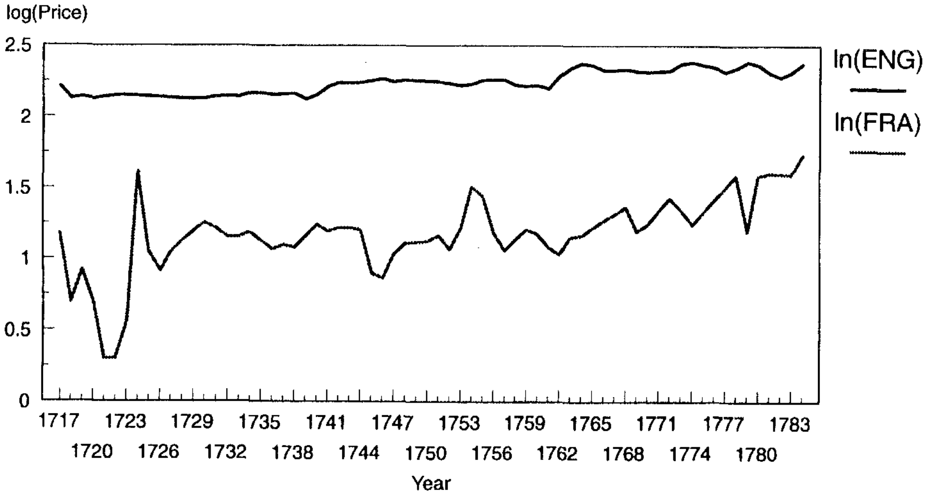

Figure 2.4. Log butter prices, 1717–1784. Price in grams of silver per kilogram of butter.

times higher in England than in France, though this ratio drops to 1.8 by the end of the period. (One possible explanation is that France's industrialization during the mid and latter parts of the 18th century drove up nontraded prices relative to England, where by 1717 the industrial revolution was already in full swing.) In contrast to wheat and charcoal, the (silver) price of butter trended upward during the period. If one thinks of dairy products as being more labor intensive than grains and charcoal, this fact is consistent with the Baumol–Bowen hypothesis, discussed in Section 3. Indeed, the price of butter relative to silver has continued to rise and stands today at roughly five times the price

Table 2.2
Augmented Dickey–Fuller regressions on English/French relative prices

$$q(i) = s - p(i) + p^*(i)$$
$$q_t(i) = \alpha_0 + \alpha_1 q_{t-1}(i) + \phi_1 (1 - L) q_{t-1}(i) + \epsilon_t(i)$$

Sample Period	Wheat 1632–1789	Charcoal 1702–1789	Butter 1719–1784
N	158	74	41
α_1	0.445	0.303	0.737
t-stat	−7.09	−8.95	−2.05
1% critical value	−3.47	−3.52	−3.60
ϕ_1	0.103	0.10	0.08
R^2	0.262	0.533	0.112
DW	2.02	1.43	1.75
σ_q	0.293	0.119	0.172

at the beginning of the eighteenth century. Put differently, if the relative price of butter to silver had remained constant since the early 17th century, butter would today cost approximately $0.43 per pound.

The third striking feature of the relative prices – especially wheat – is the strong appearance of stationarity. To check this more formally, we ran unit root tests on the English/French relative price of wheat, charcoal and butter. The results reported in Table 2.2 above are based on the augmented Dickey–Fuller specification (2.6), except that the time trend has been omitted (including a time trend does not affect the results). For wheat and charcoal, one can reject the unit root null at the 1 percent level. The autoregressive coefficient for the relative price of wheat is 0.445, implying a half life for PPP deviations of roughly one year. For charcoal, the autoregressive coefficient is only 0.303, implying a half life of only seven months.

Interestingly, we cannot reject the random-walk hypothesis for the relative price of butter, despite the relatively long time span. The estimated autoregressive coefficient is 0.74, an implied half-life of 2 years. This result that the half life of PPP deviations for butter is greater than for other commodities accords with our intuition, as butter is more likely to be nontraded than wheat or charcoal over this period. Of course, our calculations in section 2.3.3 suggest that seventy-five years may not be enough data to reliably reject nonstationarity.

3. Structural models of deviations from PPP

Until now, all the evidence we have looked at has been based on price and exchange rate data. In this section, we discuss a number of studies that attempt to explain empirically deviations from PPP in terms of more fundamental factors such as productivity, government spending, and strategic pricing decisions by firms. Our focus is on medium- to long-term movements; Frankel and Rose deal with short-term fluctuations elsewhere in this volume. We begin by reviewing some of the key theoretical issues, and then turn to the empirical evidence.

3.1. Productivity, government spending and the relative price of nontradeables

Of the many models that have been put forth to explain long-term deviations in consumption-based PPP, the most popular and enduring one is due to Balassa (1964) and Samuelson (1964). They posited that, after adjusting for exchange rates, CPIs in rich countries will be high relative to those in poor countries, and that CPIs in fast-growing countries will rise relative to CPIs in slow-growing countries. We will formalize their analysis shortly, but the main idea is as follows.

Balassa and Samuelson argue that technological progress has historically been faster in the traded goods sector than in the nontraded goods sector (perhaps because traded goods are weighted towards high-innovation agricultural or manufacturing goods) *and*, crucially, that this traded-goods productivity bias is more pronounced in high-income countries. As a consequence, CPI levels tend to be higher in wealthy countries. Why? A rise in productivity in the traded goods sector will bid up wages in the entire economy; producers of nontraded goods will only be able to meet the higher wages if there is a rise in the relative price of nontraded goods.

To take an example, consider the fact that nontraded goods are cheaper in India than in Switzerland. Although Switzerland's absolute level of productivity is higher than that of India, the productivity in its nontraded-goods sector relative to its traded-goods sector is lower.

It is important to distinguish the Balassa–Samuelson effect from the related "Baumol–Bowen" effect. Baumol and Bowen (1966) argued that within a country, there is a broad tendency for the prices of service intensive goods (education, health care, auto repair, banking, etc.) to rise over time. Historically, productivity growth in services has tended to be much slower than in more capital intensive manufacturing industries. This argument is obviously closely parallel to a key building block of the Balassa–Samuelson model, since there is a heavy overlap between nontradeables and service-intensive goods. Note, however, that the presence of a Baumol–Bowen effect is not necessarily sufficient to imply a Balassa–Samuelson effect.[30]

[30] Will there ever be a service productivity revolution to turn the Baumol–Bowen effect on its head? The possibility cannot be dismissed. Banking services have become vastly more efficient in recent years, with innovations ranging from ATMs to derivative securities; it seems plausible that the information revolution may someday lead to a long pause, if not a reversal, of the trend differential in productivity growth.

It is arguable whether one should expect to detect a Balassa–Samuelson effect in really long-run data. Even though technology can differ across countries for extended periods, the free flow of ideas together with human and physical capital produces a tendency towards long-run convergence of incomes. Of course, in the final analysis, the effect of income growth on PPP is an empirical question. Before looking at the empirical evidence, however, we examine more closely the theoretical underpinnings of the model; readers whose main interest is in the empirical material may wish to skip to the next subsection.

3.2. *A small country model of the Balassa–Samuelson effect*

In this section, we derive the central equation of the Balassa–Samuelson relationship between the real (CPI) exchange rate and the productivity differential between traded and nontraded goods.[31] One important point, generally overlooked in the literature, is that even balanced growth across the two sectors can lead to a rise in the relative price of nontradeables if nontraded goods are relatively labor intensive.

Consider the case of a small, open economy, that produces both traded and nontraded goods. The sectoral production functions are

$$Y_t^{\mathrm{T}} = A_t^{\mathrm{T}}(L_t^{\mathrm{T}})^{\theta^{\mathrm{T}}}(K_t^{\mathrm{T}})^{1-\theta^{\mathrm{T}}} \tag{3.1}$$

$$Y_t^{\mathrm{N}} = A_t^{\mathrm{N}}(L_t^{\mathrm{N}})^{\theta^{\mathrm{N}}}(K_t^{\mathrm{N}})^{1-\theta^{\mathrm{N}}} \tag{3.2}$$

where Y^{T} (Y^{N}) denote domestic output of the traded and nontraded good respectively, and K^I, L^I, and A^I are capital, labor and productivity in sector I. Let us initially assume that capital is mobile both internationally and across the two sectors internally. Assuming perfect competition in both sectors, profit maximization implies

$$R = (1-\theta^{\mathrm{T}})A^{\mathrm{T}}(K^{\mathrm{T}}/L^{\mathrm{T}})^{-\theta^{\mathrm{T}}} \tag{3.3}$$

$$R = P^{\mathrm{N}}(1-\theta^{\mathrm{N}})A^{\mathrm{N}}(K^{\mathrm{N}}/L^{\mathrm{N}})^{-\theta^{\mathrm{N}}} \tag{3.4}$$

$$W = \theta^{\mathrm{T}}A^{\mathrm{T}}(K^{\mathrm{T}}/L^{\mathrm{T}})^{1-\theta^{\mathrm{T}}} \tag{3.5}$$

$$W = P^{\mathrm{N}}\theta^{\mathrm{N}}A^{\mathrm{N}}(K^{\mathrm{N}}/L^{\mathrm{N}})^{1-\theta^{\mathrm{N}}} \tag{3.6}$$

where R is the rental rate on capital (determined in world markets), W is the wage rate (measured in tradeables) and P^{N} is the relative price of nontradeables. Since we assume no adjustment costs, it is convenient to omit time subscripts.

[31] The analysis here is based on Froot and Rogoff (1991b); [see also Rogoff (1992)].

The key result is that with perfect capital mobility, the relative price of nontradeables P^{N} is governed entirely by the production side of the economy. Equations (3.3)–(3.6) involve four equations in four variables, $K^{\mathrm{T}}/L^{\mathrm{T}}$, $K^{\mathrm{N}}/L^{\mathrm{N}}$, W, and P, which can be solved recursively as follows: Given the constant returns to scale production functions (3.1) and (3.2), eq. (3.3) implies a unique level of $K^{\mathrm{T}}/L^{\mathrm{T}}$ consistent with the world rate of return on capital R. Given $K^{\mathrm{T}}/L^{\mathrm{T}}$, eq. (3.5) determines the economy-wide wage rate W. The remaining two equations (3.4) and (3.6) then determine $K^{\mathrm{N}}/L^{\mathrm{N}}$ and P.

By log-differentiating eqs.(3.3)–(3.6), one can obtain a (slight) generalization of the classic Balassa–Samuelson hypothesis:

$$\widehat{p} = (\theta^{\mathrm{N}}/\theta^{\mathrm{T}})\widehat{a}^{\mathrm{T}} - \widehat{a}^{\mathrm{N}} \tag{3.7}$$

where $\widehat{x} \equiv d \log x$. If both sectors have the same degree of capital intensity – if $\theta^{\mathrm{N}} = \theta^{\mathrm{T}}$ – then the percentage change in the relative price of traded goods is simply equal to $\widehat{a}^{\mathrm{T}} - \widehat{a}^{\mathrm{N}}$, the productivity growth differential between the traded and nontraded sectors. If, however $\theta^{\mathrm{N}} > \theta^{\mathrm{T}}$ (one generally thinks of nontraded goods as being more labor intensive), then even balanced productivity growth ($\widehat{a}^{\mathrm{T}} = \widehat{a}^{\mathrm{N}}$) will lead to an appreciation of the relative price of traded goods.[32]

Note that in the small open economy with perfect factor mobility, demand factors do not affect P^{N}, they only affect a country's consumption basket.[33] This is not the case for an economy fully (or partially) shut off from world capital markets, since R is no longer tied down by world markets. (The same is true, of course, for a large economy.) In this case, eqs. (3.3)–(3.6) must be supplemented by the demand side of the model.[34]

Demand factors can also be important in the small-country case *in the short run* if labor and/or capital cannot be transferred instantly across sectors.[35] Froot and Rogoff (1991) show that in this case, government spending will tend to raise the relative price of nontradeables, if the spending falls disproportionately on nontraded goods, relative to private expenditure shares. Rogoff (1992) shows that the model also implies that temporary shocks to traded goods productivity can have *permanent* effects on P^{N}. The reason is that private agents can use international capital markets to smooth their consumption of traded goods. As a result the relative *intratemporal* price of traded and nontraded goods is smoothed.

[32] De Gregorio and Wolf (1994) extend this model to allow for changes in the terms of trade (the relative price of importables and exportables).

[33] For further discussion, see Obstfeld and Rogoff (1995), ch. 4.

[34] This is a short-run result. In a representative agent model with constant discount rate β, the long-run interest rate is again tied down, and other demand factors do not enter into the determination of P^{N}.

[35] Demand factors will matter in the long run provided there is some fixed factor (e.g. land) which can be transferred across sectors.

Having discussed the basic theory underpinning the Balassa–Samuelson approach, we are now ready to examine the empirical literature.

3.3. Long-term productivity differentials and the real exchange rate

Balassa was the first to formally test the proposition that richer countries have higher real exchange rates; Balassa (1964), for example, reports the following regression for a cross-section of twelve industrial countries for the year 1960:[36]

$$\begin{aligned}(P/SP^*)_i = \underset{0.49}{\alpha} + \underset{0.51\ (8.33)}{\beta(\mathrm{GNP/POP})_i}\end{aligned} \tag{3.8}$$

where P/SP^* is the inverse of the level of the real exchange rate, GNP/POP is GNP/ population, and t statistics are in parentheses; the regression has 10 degrees of freedom. Balassa (1973) presents similar results, again finding that richer countries have higher (exchange-rate adjusted) price levels. Officer (1976b) surveys a host of follow-up studies that, on the whole, yielded much more negative results. Officer argued that Balassa's results are extremely sensitive to the year chosen and to the countries included in the regression. Note that eq. (3.8) is a test of *absolute* purchasing power parity; some of the data sources used for early absolute PPP comparison include Gilbert and Kravis (1954), Gilbert et al. (1958), and Kravis et al. (1975).

The most recent effort to construct absolute comparisons of PPP is Summers and Heston (1991), who construct absolute PPP data for a broad range of countries. Generally, their data reveal striking differences in price levels between poor countries as a group and rich countries as a group. Once divided into two groups, the within-group correlations between income and price level are much less apparent (see their figure on p. 336).

The preceding studies dealt with the cross-sectional implications of Balassa–Samuelson. Hsieh (1982) was the first to look at time series implications. His study focused on Japanese and German real exchange rates vis-à-vis the United States for the years 1954–1976. Hsieh's central regressions were of the form:

$$\begin{aligned}p_t - s_t - p_t^* = \beta_0 + \beta_1\left[a_t^{\mathrm{T}} - a_t^{\mathrm{N}}\right] - \beta_2\left[a_t^{*\mathrm{T}} - a_t^{*\mathrm{N}}\right]\\ +\beta_3\left[w_t - s_t - w_t^* + a_t^{\mathrm{T}} - a_t^{*\mathrm{T}}\right]\end{aligned} \tag{3.9}$$

where w is the (log) nominal wage rate. Hsieh found that the productivity differential variables were significant and of the correct sign for both real exchange rates, and that his OLS regression results were robust both to correcting for serial correlation and to using instrumental variables techniques. It should be

[36] Balassa does not provide the standard error on α or R^2; the correlation coefficient is 0.92.

noted that the variable $[w_t - s_t - w_t^* + a_t^{\mathrm{T}} - a_t^{*\mathrm{T}}]$, which includes nominal wage differentials, is extremely highly correlated with the lagged real exchange rate. Thus Hsieh's results may be sensitive to whether or not this "error correction" term is included.

Marston (1987) and Edison and Klovland (1987) also present evidence suggestive of a Balassa–Samuelson effect. Edison and Klovland examine time series data on the real exchange between the British pound and the Norwegian krone for the years 1874–1971. This long time period allows them to detect significant evidence of a productivity differential effect using as proxies both the real output differential and a measure of the commodity/service productivity ratio differential.

Marston (1987) looks at the yen/dollar real exchange rate over the period 1973–1983, and calculates traded-nontraded goods productivity differentials using OECD data that disaggregates the economy into ten subsectors. It is worthwhile to digress briefly to explain his aggregation approach. He designates two sectors as traded: manufacturing; and agricultural, hunting, fishing and forestry. Six sectors are deemed nontraded: construction; wholesale and retail trade; restaurants and hotels; transport, storage and communication; finance, insurance, real estate; business services, community, social and personal services; and government services. Marston excludes the mining and quarrying sector, because it is energy intensive and therefore was very sensitive to OPEC pricing policies over the period. For the same reason, he excludes the electricity, gas and water sector.

Using sectoral employment data, Marston calculates labor productivity differentials between traded and nontraded goods, and argues that these variables provide an extremely plausible explanation of the long-run trend real appreciation of the yen against the dollar.

The evidence of later studies is somewhat mixed. Froot and Rogoff (1991a,b) look at a cross-section of 22 OECD countries for the years 1950–1989. They find that the correlation between productivity differentials and the real exchange rate is weak at best, both for their full sample and for various subsamples.

Asea and Mendoza (1994), take a different approach; their analysis is based on a dynamic two-country general equilibrium model. They take sectoral OECD data to calculate relative traded goods prices for fourteen OECD countries over the period 1975–1985. They first regress the relative price of nontraded goods for each country against traded-nontraded productivity differentials, and then regress cross-country real exchange rates against the relative price of nontraded goods (they try both actual and estimated). Asea and Mendoza conclude that although the productivity differentials between traded and nontraded goods are extremely significant in explaining changes in the relative price of nontraded goods within each country, changes in nontraded goods prices account for only a small and insignificant part of real exchange-rate changes across countries (using either CPI or GDP deflators). Thus while the data reveal evidence of a Baumol–Bowen effect, the Balassa–Samuelson effect is more difficult to detect.

De Gregorio, Giovannini, and Wolf (1994a) similarly conclude that differences in productivity growth across traded and nontraded goods help explain the relative price of nontraded goods. (However, they reach somewhat more positive conclusions than Asea and Mendoza concerning the ability of productivity differentials to explain changes across countries in the real exchange rate; we will discuss their results in more detail shortly.)

Note that if two countries have different weights on services in their consumption baskets, then the Baumol–Bowen effect alone can produce significant trend movements in CPI real exchange rates, even in the absence of a Balassa–Samuelson effect. Suppose, for example, that two countries have identical technologies at all times, but that one country has a higher share of services in the CPI. Then the presence of a Baumol–Bowen effect is sufficient to yield a trend in the CPI real exchange rate. If the Baumol–Bowen effect is indeed important – the evidence in both Asea and Mendoza, and in De Gregorio, Giovannini, and Wolf strongly suggests that it is – then one must have convergence in *tastes*, not just technologies, for the real exchange rate to converge in the long run.

3.4. Demand factors and the real exchange rate

A striking feature of the Balassa–Samuelson model developed in Section 3.2 is that the real exchange rate depends entirely on supply factors; demand factors do not enter. This property of the model depends on several assumptions:

(1) the country is small and cannot affect the world interest rate;

(2) capital is mobile internationally;

(3) both capital and labor are instantaneously mobile across sectors internally;

(4) there are constant returns to scale in the mobile factors (i.e. there is no third factor in production such as land which is immobile across sectors).

If, for example, capital and labor are mobile across sectors in the long run but not in the short run, demand factors can have a short-run impact on the real exchange rate. The possibility that demand factors may matter, at least in the short run, has been explored empirically by Froot and Rogoff (1991a,b), Rogoff (1992) and by De Gregorio, Giovannini, and Wolf (1994a).[37]

Froot and Rogoff (1991a) look at alternative explanations for the significant shifts in real exchange rates over the EMS period. Between 1986 and 1991,

[37] Ahmed (1987) also looks empirically at the effects of government spending on the real exchange rate. His model, however, focuses on home-produced versus foreign goods (the terms of trade), as opposed to traded versus nontraded. He assumes that government spending falls more heavily on foreign goods than does spending by domestic consumers, and he finds support for this hypothesis on historical data for Great Britain.

for example, Italy's CPI inflation rate exceeded Germany's by more than 15 percent, while the lira/mark exchange rate remained fixed. Froot and Rogoff explore to what extent relative growth in Italian government spending might account for this phenomenon. In their model, it is assumed that government spending falls disproportionately on nontraded goods, thereby bidding up their relative price. They regress the real CPI exchange rate against various measures of productivity differentials and government spending (as a ratio to GNP). The government spending variable consistently enters with correct sign in all the individual country regressions and is strongly significant in the pooled time series cross-section regressions. Neither productivity differentials or government spending enters significantly.

Froot and Rogoff suggest that government spending effects, because they are transitory, may be difficult to pick up in highly volatile floating exchange rate data. To pursue this conjecture, Froot and Rogoff (1991b) look at data for twenty-two OECD countries for the period 1950–1989, a sample which includes both fixed and floating rate data. They also modify their earlier model to allow for gradual factor adjustment across sectors, implying that the effects of government spending will only be temporary. Overall, they find that the government spending differential consistently enters pooled regressions significantly and with the correct sign, both over the full sample and over separate fixed and floating rate periods. Though factor mobility causes the effect to die out over the long run, the half life appears quite long – more than five years.

Rogoff (1992) estimates a related model on quarterly data for the yen/dollar rate over the period 1975–1990. Although government spending appears to be highly correlated with the real exchange rate, it does not enter significantly into the regressions once one controls for shocks to the world price of oil.

De Gregorio, Giovannini, and Wolf (1994a) present a cross-country panel regression that attempts to sort of the importance of demand and supply factors. Their model is closely related to the model presented in Section 3.1. Like Asea and Mendoza (1994), they use the OECD intersectoral data base to construct measures of productivity growth in the traded and nontraded goods sectors. The version they use covers fourteen countries[38] and twenty sectors; the data include both real and nominal output permitting construction of sectoral price deflators as well as detailed input data allowing derivation of total factor productivity levels.[39] Their method for classifying sectors between traded and nontraded is somewhat different than Marston's, though yielding broadly similar results. De Gregorio, Giovannini, and Wolf calculate total exports of each sector across all fourteen OECD countries and take the ratio to total production. They define

[38] Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.

[39] Stockman and Tesar (1995) have previously used this same data set in their work on international real business cycles.

a sector as tradeable if more than ten percent is exported. This leads them to classify as tradeables all of manufacturing, agriculture and mining; they also classify transportation services as tradeable. All other services, comprising 50–60 percent of GNP are classified as nontradeable. De Gregorio, Giovannini, and Wolf then calculate total factor productivity using Solow residuals.[40] They also test for the effect of government spending on the relative price of nontradeables.

Their results are very interesting and instructive. De Gregorio, Giovannini, and Wolf's central regression is of the form:[41]

$$\left(p^{\mathrm{N}}-p^{\mathrm{T}}\right)_{i,t}=\alpha_i+\beta_1\left[\left(\theta^{\mathrm{N}}/\theta^{\mathrm{T}}\right)a^{\mathrm{T}}-a^{\mathrm{N}}\right]_{i,t}+\beta_2 g_{i,t}+\beta_3 y_{i,t} \tag{3.10}$$

where g is real government spending (excluding government investment) over real GDP, y is real per capita income, i subscripts denote country and t subscripts denote time. Note that the weight on tradeable goods productivity growth a^{T} is greater than the weight on a^{N} if $\theta^{\mathrm{N}}>\theta^{\mathrm{T}}$, which is very plausible. As we discussed in Section 3.1, balanced productivity growth is likely to raise the relative price of nontradeables in the small open economy. Because the rate of return on capital is tied down by international capital mobility, productivity growth raises the wage/rental ratio, and therefore raises the relative price of the labor-intensive good.[42] In constructing g, the share of government spending in GNP, De Gregorio, Giovannini, and Wolf use separate deflators to convert nominal government spending to nominal GDP.[43] Otherwise, because government spending has a higher share of nontraded goods than overall GDP, changes in the relative price of nontradeables will affect the ratio even if quantities are constant. (The OECD data permits this adjustment.)

Pooling data for all fourteen countries over the full 1971–1985 sample period (they have 210 observations), De Gregorio, Giovannini, and Wolf obtain

$$\begin{array}{lllll}\left(p^{\mathrm{N}}-p^{\mathrm{T}}\right)_{i,t}=\alpha_i+ & \beta_1\left[\left(\theta^{\mathrm{N}}/\theta^{\mathrm{T}}\right)a^{\mathrm{T}}-a^{\mathrm{N}}\right]_{i,t}+ & \beta_2 g_{i,t} & + & \beta_3 y_{i,t}\\ & 0.234 & 1.974 & & 0.281\\ & (0.018) & (0.119) & & (0.030)\end{array}$$

[40] De Gregorio et al. use of total factor productivity, which controls for capital inputs, contrasts with Marston's use of labor productivity. (Asea and Mendoza also use total factor productivity.) It is not obvious, however, that the total factor productivity is necessarily superior, since data on capital inputs is notoriously unreliable.

[41] They also consider a variant of their empirical model where the lagged rate of change of nontraded goods prices enters into the regression; this modification does not substantially affect the results.

[42] Strictly speaking, the coefficient in eq. (3.10) is correct for the case of perfect factor mobility. When factors are immobile across sectors, the weights depend on shares in aggregate demand; [see Rogoff (1992)].

[43] Froot and Rogoff (1991b) also attempt to control for this effect by using CPI and WPI data to construct separate deflators for government spending and GNP, but their data are much cruder.

where standard errors are in parentheses. The productivity, government spending and income variables are all highly significant and of the theoretically predicted signs.

Although their model is not explicitly dynamic, De Gregorio, Giovannini, and Wolf attempt to see whether demand factors matter in the long run by averaging data for each country over time, and running a regression for the cross-section data. They find that over the long run, the productivity differentials remain extremely significant whereas the effects of demand factors (government spending and income) become less important. It would be interesting to explore this issue further by estimating a model that explicitly accounts for dynamic adjustment. (Asea and Mendoza (1994) and Rogers (1994) both represent useful efforts in this direction.)

Recently, De Gregorio and Wolf (1994) have extended this analysis to incorporate term of trade shocks (shocks to the relative price of home exports versus home imports). They find that the terms of trade are important empirically, though productivity and government spending differentials continue to be important. Relative incomes, however, become insignificant when terms of trade shocks are included. They conclude that the income variable in the above regression may be proxying for terms of trade shocks.

3.5. Pricing to market

Most of our discussion in this section has focused on deviations from the law of one price due to the presence of nontraded goods. We have paid relatively little attention to the factors that may cause deviations from the law of one price in traded goods (except to note that one must be careful to recognize that many goods that appear to be highly "traded" in fact have a large nontraded component). The empirical evidence, including Isard (1977), Giovannini (1988) and Engel (1993), strongly suggests that deviations from the law of one price in traded goods are important in practice and that the short-run size and direction of these departures appears to track closely nominal exchange rate movements. One obvious explanation is that of short-term price rigidities, due, say, to menu costs in changing prices. Frankel and Rose, as well as Garber and Svensson deal with price rigidities elsewhere in this volume, discussing how short-term nominal rigidities can affect the transmission of real and monetary disturbances.

Another theory of why there can be deviations from the law of one price in traded goods is the "pricing to market" theory of Krugman (1987) and Dornbusch (1987). This literature is also covered by Feenstra elsewhere in this Handbook, so our discussion here is brief. In the pricing to market framework, oligopolistic suppliers are able to charge different prices for the same good in different countries; thus the prices of BMWs can differ between Germany and the United States. If the BMWs are truly tradeable, why isn't this gap closed by goods market arbitrage? The theory posits that there are important cases where companies can separately license the sale of goods at home and abroad. Of course, the ability to price discriminate across markets may be very limited in practice. In electronics, for example, there exists an active "gray market" in which goods are purchased in low-price countries for immediate resale in countries where the manufacturer is attempting to charge a higher price.[44]

In addition to potentially explaining longer-run deviations from the law of one price, "pricing to market" theories also have implications for the transmission of monetary disturbances if there are nominal rigidities. The Dornbusch–Krugman models assume that in the short-run, costs are set in nominal terms in the currency of the supplier. Then, if there is an exogenous appreciation in the home country's nominal exchange rate (the pricing to market literature is partial equilibrium), the real cost of supplying goods for foreign sale rises. If demand were unit elastic, the markup of price over cost would not be affected [see Marston, (1990)], but the markup over cost on foreign goods will fall if the foreign price elasticity of demand is greater than unity. (This would generally be the case for a monopolist with non-zero marginal cost of production.) Indeed, much of the empirical literature focuses on short-run transmission effects.

Kasa (1992) questions the price discrimination story as the underlying rationale for pricing to market. Why do studies such as Knetter (1989) find relatively similar effects of exchange rates on markups across industries, if price discrimination is central?[45] Kasa argues that pricing to market is better rationalized using an adjustment cost framework, where either the firm faces some kind of adjustment cost in changing prices or, as in Froot and Klemperer (1989), consumers face fixed costs in switching between products. Froot and Klemperer point out that if their adjustment cost story is correct, then changes in exchange rates that are expected to be temporary should lead to much greater fluctuations in markups than deviations that are expected to be permanent. They present some evidence that is suggestive of their theory, under the assumption that private agents generally viewed the mid-1980s run up in the dollar as temporary.

A number of studies, including Knetter (1989), find that pricing to market is more pronounced for German and Japanese exporters than it is for American exporters. Recently, some explanations for this stylized fact have been unearthed. Knetter (1993) finds that pricing to market behavior seems more similar across countries if one controls for industry effects; U.S. exporters tend to be concentrated in industries where, globally, pricing to market behavior is less pronounced. Rangan and Lawrence (1993) argue that part of the reason

[44]The main defense against the "gray market" is for manufacturers to refuse to honor the warranty except in the original country of sale. This strategy is obviously more likely to be successful in the case of autos than for, say, VCRs.

[45]Elsewhere in this volume, Feenstra argues that the differences in "pass-through" of exchange-rate changes across industries are in fact quite large.

U.S. pricing to market behavior seems less pronounced is that many U.S. firms sell their products abroad to subsidiaries, and that the pricing to market behavior takes place at the subsidiary level.

Ghosh and Wolf (1994) try to discriminate between menu costs and pricing to market theories. Their data set consists of cover prices of the magazine *The Economist* for eleven European countries and the United States for the years 1973–1990. They argue that importance of lagged exchange rate changes on relative price changes (across countries) supports the view that deviations from the law of one price must be driven at least partly by menu costs. As we have argued above, however, the pricing to market and menu cost theories of PPP deviations are not mutually exclusive.

Finally, in an interesting recent paper, Feenstra and Kendall (1994) argue that changes in price markups across countries over time may have a permanent component and that, empirically, this effect may in some cases be as important as the Balassa–Samuelson effect in explaining deviations from the law of one price. The existence of a permanent component, of course, essentially requires that firms be able to maintain segmentation across the markets indefinitely.

4. Conclusions

Over the past ten years, research on purchasing power parity has enjoyed a rebirth, partly due to innovations in econometrics, and even more to the development of new data sets that allow researchers to investigate both longer and more disaggregated time series. The main positive result is that there does seem to be long-run convergence to PPP, though further work on the issue of survivorship bias would be valuable. Also, the most convincing evidence on long-run convergence to PPP still comes from data sets that employ at least some fixed-rate data. Perhaps by the time of the next Handbook, there will be more years of floating-rate data from more countries and perhaps, if these are combined with more powerful econometric techniques, we will have a clearer picture of whether and how fast exchange rates converge to PPP under floating rates.[46]

There has also been a considerable amount of progress in recent years in analyzing the effect of productivity and government spending shocks on real exchange rates. Most of the empirical literature, however, does not explicitly take dynamics into account. As we have seen in the simple PPP literature, dynamics can be quite central, so further progress on understanding the dynamic effects of various real shocks might potentially prove fruitful.

[46]Since this survey was written, Frankel and Rose (1995), and Wei and Parsley (1995), have attempted to demonstrate mean reversion using only post-1973 data. Wei and Parsley look at tradeable goods prices for a cross-section of 14 OECD countries. Frankel and Rose use CPI data for a broader cross-section, including some high-inflation countries. Interestingly, both studies find half lives for PPP deviations very close to the 4-year consensus estimate noted here.

References

Abuaf, N. and P. Jorion (1990), "Purchasing power parity in the long run", Journal of Finance 45:157–174.

Adler, M. and B. Lehman (1983), "Deviations from purchasing power parity in the long run", Journal of Finance 39:1471–1487.

Ahmed, S. (1987), "Government spending, the balance of trade and terms of trade in British history", Journal of Monetary Economics 20:195–220.

Asea, P.K. and E. Mendoza (1994), "The Balassa–Samuelson model: A general equilibrium appraisal", Review of International Economics, 2:244–267.

Balassa, B. (1964), "The purchasing power parity doctrine: A reappraisal", The Journal of Political Economy 72:584–596.

Balassa, B. (1973), "Just how misleading are official exchange rate conversions: A comment", Economic Journal 83:1258–1267.

Banerjee, A., J. Dolado, D. Hendry, and G. Smith (1986), "Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence", Oxford Bulletin of Economics and Statistics 48:253–277.

Baumol, W. (1993), "Social wants and dismal science: The curious case of climbing costs on health and teaching", mimeo, Princeton University.

Baumol, W. and W. Bowen (1966), Performing arts: The economic dilemma (The Twentieth Century Fund, New York).

Bayoumi, T., P. Clark, S. Symansky, and M. Taylor (1994), "Robustness of equilibrium exchange rate calculations to alternative assumptions and methodologies", in: J. Williamson, ed., Estimating equilibrium exchange rates (Institute for International Economics, Washington DC).

Beveridge, W. (1939), Prices and wages in England: From the twelfth to the nineteenth century, vol. 1 (Longman, Greens and Co., London).

Breuer, J. (1994), "Purchasing power parity: A survey of and challenge to recent literature", in: J. Williamson, ed., Estimating equilibrium exchange rates (Institute for International Economics, Washington DC).

Bryant, M. and S. Cecchetti (1993), "The consumer price index as a measure of inflation", Economic Review of the Federal Reserve Bank of Cleveland Q4:15–24.

Campbell, J. and R. Clarida (1987), "The dollar and real interest rates: An empirical investigation", in: K. Brunner and A. Meltzer, eds., Carnegie–Rochester Series on Public Policy 27:103–140.

Campbell, J. and P. Perron (1991), "Pitfalls and opportunities: What every macroeconomist should know about unit roots", in: S. Fischer and O. Blanchard, eds., NBER macroeconomics annual (MIT Press, Cambridge, MA).

Cassel, G. (1921), The world's money problems (E.P. Dutton and Co., New York).

Cassel, G. (1922), Money and foreign exchange after 1914 (Macmillan, New York).

Cavallo, D. (1986), "Estadisticas de la evolucion economica de Argentina, 1913–1984", Estudios 9:103–184.

Cheung, Y. and K. Lai (1993a), "Long-run purchasing power parity during the recent float", Journal of International Economics 34:181–192.

Cheung, Y. and K. Lai (1993b), "A fractional cointegration analysis of purchasing power parity", Journal of Business and Economic Statistics 11:103–112.

Cheung, Y. and K. Lai (1993c), "Mean reversion in real exchange rates", mimeo, UCLA.

Chowdhury, A. and F. Sdogati (1993), "Purchasing power parity in the major EMS countries: The role of price and exchange rate adjustment", Journal of Macroeconomics 15:25–45.

Christiano, L. (1992), "Searching for a break in GNP", Journal of Business and Economic Statistics 10:237–500.

Corbae, D. and S. Ouliaris (1988), "Cointegration and tests of purchasing power parity", Review of Economics and Statistics 70:508–521.

Cumby, R. (1993), "Forecasting exchange rates on the hamburger standard: what you see is what you get with McParity", mimeo, Stern School of Business.

Cumby, R. and J. Huizinga (1991), "The predictability of real exchange rate changes in the short and long run", Japan and the World Economy 3:17–38.

Darby, M. (1983), "Movements in purchasing power parity: The short and long runs", in: M. Darby and J. Lothian, eds., The international transmission of inflation (University of Chicago Press, Chicago).
Davutyan, N. and J. Pippenger (1985), "Purchasing power parity did not collapse during the 1970's", American Economic Review 75:1151–1158.
Diebold, F., S. Husted, and M. Rush (1991), "Real exchange rates under the gold standard", Journal of Political Economy 99:1252–1271.
De Gregorio, J., A. Giovannini, and H. Wolf (1994a), "International evidence on tradables and nontradables inflation", European Economic Review, 38:1225–1244.
De Gregorio, J., A. Giovannini and H. Wolf (1994b), "The behavior of nontradeable goods prices in Europe: Evidence and interpretation", Review of International Economics, 2:284–305.
De Gregorio, J., and H. Wolf (1994), "Terms of trade, productivity, and the real exchange rate", National Bureau of Economic Research Working Paper No. 4807.
Dickey, D. and W. Fuller (1979), "Distribution of the estimators for autoregressive time series with a unit root", Journal of the American Statistical Association 74:427–431.
Dornbusch, R. (1976), "Expectations and exchange rate dynamics", Journal of Political Economy 84:1161–1176.
Dornbusch, R. (1987), "Purchasing power parity", in: J. Eatwell, M. Migare, and P. Newman, eds., The New Palgrave Dictionary (Stockton Press, New York).
Edison, H. (1987), "Purchasing power parity in the long run: A test of the dollar/pound exchange rate (1890–78)", Journal of Money, Credit, and Banking 19:376–387.
Edison, H., J. Gagnon, and W. Melick (1994), Understanding the empirical literature on purchasing power parity: the post-Bretton Woods era (Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC).
Edison, H. and J.T. Klovland (1987), "A quantitative reassessment of the purchasing power parity hypothesis: Evidence from Norway and the United Kingdom", Journal of Applied Econometrics 2:309–333.
Elliot, G., Rothenberg, T., and J. Stock (1992), "Efficient tests for an autoregressive root", National Bureau of Economic Research Technical Working Paper No. 130.
Enders, W. (1988), "Arima and cointegration tests of PPP under fixed and flexible exchange rate regimes", Review of Economics and Statistics 70:504–508.
Engel, C. (1993), "Is real exchange rate variability caused by relative price changes? An empirical investigation", Journal of Monetary Economics 32:35–50.
Engel, C. and J. Rogers (1994), "How wide is the border?", National Bureau of Economic Research Working Paper No. 4829.
Engle, R. and C. Granger (1987), "Cointegration and error correction: Representation, estimation, and testing", Econometrica 55:251–276.
Feenstra, R., and J. Kendall (1994), "Pass-through of exchange rates and purchasing power parity", National Bureau of Economic Research Working Paper No. 4842.
Feldstein, M. (1993), "The dollar and the trade deficit", National Bureau of Economic Research Working Paper No. 4325.
Fisher, E. and J. Park (1991), "Testing purchasing power parity under the null hypothesis of co-integration", The Economic Journal 101:1476–1484.
Flynn, N. and J. Boucher (1993), "Tests of long run purchasing power parity using alternative methodologies", Journal of Macroeconomics 15:109–122.
Frankel, J. (1986), "International capital mobility and crowding out in the U.S. economy: Imperfect integration of financial markets or goods markets?", in: R. Hafer, ed., How open is the US economy (Lexington Books, Lexington).
Frankel, J. (1990), "Zen and the art of modern macroeconomics: The search for perfect nothingness", in: W. Haraf and T. Willett, eds., Monetary policy for a volatile global economy. (American Enterprise Institute, Washington, DC).
Frankel, J. and A. Rose (1995), "A panel project on purchasing power parity: Mean reversion within and between countries", National Bureau of Economic Research Working Paper No. 5006.
Frenkel, J. (1978), "Quantifying international capital mobility in the 1980s", in: D. Bernheim and J. Shoven, eds., National saving and economic performance (University of Chicago Press, Chicago) 227–260.

Frenkel, J. (1981), "The collapse of purchasing power parity during the 1970s", European Economic Review 16:145–165.

Friedman, M. (1957), "The case for flexible exchange rates", in: Essays in positive economics (University of Chicago Press, Chicago).

Froot, K. and P. Klemperer (1989), "Exchange-rate pass through when market share matters", American Economic Review 79:637–654.

Froot, K. and K. Rogoff (1991a), "The EMS, the EMU, and the transition to a common currency", in: S. Fisher and O. Blanchard, eds., National bureau of economic research macroeconomics annual (MIT Press, Cambridge, MA) 269–317.

Froot, K. and K. Rogoff (1991b), "Government consumption and the real exchange rate: The empirical evidence", mimeo, Harvard Business School.

Froot, K., Kim, M., and K. Rogoff (1995), "The law of one price over seven hundred years", mimeo, Princeton University.

Ghosh, A. and H. Wolf (1994), "Pricing in international markets: Lessons from the Economist", National Bureau of Economic Research Paper No. 4806.

Gilbert, M. and Associates (1958), Comparative national price products and price levels: A study of western Europe and the United States (Organization for European Economic Cooperation, Paris).

Gilbert, M., and I. Kravis (1954), An international comparison of national products and the purchasing power of currencies: A study of the United States, the United Kingdom, France, Germany, and Italy (Organization for European Economic Cooperation, Paris).

Giovannetti, G. (1992), "A survey of recent empirical tests of the purchasing power parity hypothesis", Banca Nazionale del Lavoro Quarterly Review 180:81–101.

Giovannini, A. (1988), "Exchange rates and traded goods prices", Journal of International Economics 24:45–68.

Glen, J. (1992), "Real exchange rates in the short, medium, and long run", Journal of International Economics 33:147–166.

Grilli, V. and G. Kaminsky (1991), "Nominal exchange rate regimes and the real exchange rate: Evidence from the United States and Great Britain, 1885–1986", Journal of Monetary Economics 27:191–212.

Hakkio, C. (1984), "A reexamination of purchasing power parity", Journal of International Economics 17:265–277.

Hakkio, C.S. and M. Rush (1991), "Cointegration: How short is the long run?", Journal of International Money and Finance 10:571–581.

Hauser, H. (1936), Recherches et document sur l'histoires des prix en France de 1500 a 1800 (Slatkive, Geneva).

Horvath, M. and M. Watson (1993), "Testing for cointegration when some of the cointegrating vectors are known", Federal Reserve Bank of Chicago Working Paper No. 93–15.

Hsieh, D. (1982), "The determination of the real exchange rate: The productivity approach", Journal of International Economics 12:355–362.

Huizinga, J. (1987), "An empirical investigation of the long-run behavior of real exchange rates", in: K. Brunner and A. Meltzer, eds., Carnegie-Rochester Series on Public Policy 27:149–215.

Isard, P. (1977), "How far can we push the law of one price?", American Economic Review 67:942–948.

Jastram, R. (1981), Silver: The restless metal (Wiley, New York).

Johansen, S. (1988), "Statistical analysis of cointegration vectors", Journal of Economic Dynamics and Control 12:231–254.

Johansen, S. (1991), "Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models", Econometrica 59:1551–1581.

Johansen, S. and K. Juselius (1992), "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for U.K.", Journal of Econometrics 53:211–244.

Johnson, D.R. (1990), "Co-integration, error correction, and purchasing power parity between Canada and the United States", Canadian Journal of Economics 23:839–855.

Johnson, P. (1991), "Aggregate price indices, cointegration, and tests of the purchasing power parity hypothesis", Economics Letters 36.

Kasa, K. (1992), "Adjustment costs and pricing to market: Theory and evidence", Journal of International Economics 32:1–30.

Keynes, J. (1932), Essays in persuasion (Harcourt Brace, New York).

Kim, H. (1990), "Purchasing power parity in the long run: A cointegration approach", Journal of Money, Credit, and Banking 22:491–503.

Knetter, M. (1989), "Price discrimination by U.S. and German exporters", American Economic Review 79:198–210.

Knetter, M. (1993), "International comparisons of pricing to market behavior", American Economic Review 83:473–486.

Kravis, Irving B. (1975), A system of international comparisons of gross product and purchasing power, Phase 1 (Johns Hopkins University Press, Baltimore).

Kravis, I. and R. Lipsey (1988), "National price levels and the prices of tradeables and nontradeables", American Economic Review 78:474–478.

Krugman, P. (1978), "Purchasing power parity and exchange rates", Journal of International Economics 8:397–407.

Krugman, P. (1987), "Pricing to market when the exchange rate changes", in: S. Arndt and J.D. Richardson, eds., Real-financial linkages among open economies (MIT Press, Cambridge, MA).

Krugman, P. (1990), "Equilibrium exchange rates", in: W. Branson, J. Frenkel, and M. Goldstein, eds., International monetary policy coordination and exchange rates (The University of Chicago Press, Chicago).

Kugler, P. and C. Lenz (1993), "Multivariate cointegration analysis and the long run validity of PPP", Review of Economics and Statistics 75:180–184.

Lastrapes, W. (1992), "Sources of fluctuations in real and nominal exchange rates", Review of Economics and Statistics 74:530–539.

Lebow, D., J. Roberts, and D. Stockton (1992), "Economic performance under price stability", Board of Governors of the Federal Reserve System Working Paper No. 125.

Lee, M. (1974), Purchasing power parity (Marcel Dekker, New York).

Lothian, J. and M. Taylor (1994), Real exchange rate behavior: The recent float from the perspective of the past two centuries (International Monetary Fund, Washington DC), Journal of Political Economy, forthcoming.

Mark, N. (1990), "Real exchange rates in the long run: An empirical investigation", Journal of International Economics 28:115–136.

Marston, R. (1987), "Real exchange rates and productivity growth in the United States and Japan", in: S. Arndt and J.D. Richardson, eds., Real-financial linkages among open economies (MIT Press, Cambridge, MA).

Marston, R. (1990), "Pricing to market in Japanese manufacturing", Journal of International Economics 29:217–236.

McKinnon, R. (1971), "Monetary theory and control flexibility in the foreign exchanges", Princeton University, Essays in International Finance 84.

McKinnon, R.I. (1984), An international standard for monetary stabilization (Institute for International Economics, Washington, DC).

McNown, R. and M. Wallace (1989), "National price levels, purchasing power parity, and cointegration: A test of four high inflation economies", Journal of International Money and Finance 8:533–546.

Meese, R. and K. Rogoff (1983), "Empirical exchange rate models of the seventies: do they fit out of sample?", Journal of International Economics 14:3–24.

Meese, R. and K. Rogoff (1988), "Was it real? The exchange rate interest differential relation over the modern floating exchange rate period", Journal of Finance 43:933–948.

Mussa, M. (1986), "Nominal exchange rate regimes and the behavior of real exchange rates: Evidence and implications", in: K. Brunner and A. Meltzer, eds., Carnegie–Rochester Series on Public Policy 25:117–214.

Obstfeld, M. and K. Rogoff (1995), "Exchange rate dynamics redux", Journal of Political Economy 102, forthcoming.

Obstfeld, M. and K. Rogoff (1996), Foundations of international macroeconomics, (MIT Press, Cambridge, MA) forthcoming.

Officer, L. (1976a), "The purchasing power parity theory of exchange rates: A review article", International Monetary Fund Staff Papers 23:1–60.
Officer, L. (1976b), "The productivity bias in purchasing power parity: An econometric investigation", International Monetary Fund Staff Papers 10:545–579.
Ogaki, M. and J. Park (1991), "Inference in cointegrated models using VAR prewhitening to estimate short-run dynamics", University of Rochester Center for Economic Research Working Paper No. 281.
Papell, D. (1994), "Exchange rates and prices: An empirical analysis", International Economic Review 35:397–410.
Perron, P. (1988), "Trends and random walks in macroeconomic time series: Further evidence from a new approach", Journal of Economic Dynamics and Control 12:277–302.
Perron, P. (1989), "The great crash, the oil price shock, and the unit root hypothesis", Econometrica 57:1361–1401.
Phillips, P. (1987a), "Time series regression with a unit root", Econometrica 55:297–332.
Phillips, P. (1987b), "Asymptotic expansions in nonstationary vector autoregressions", Econometric Theory 3:45–68.
Poterba, J. and L. Summers (1986), "The persistence and volatility of stock market fluctuations", American Economic Review 76:1142–1151.
Rangan, S. and R. Lawrence (1993), "The responses of US firms to exchange rate fluctuations: Piercing the corporate veil", Brookings Papers on Economic Activity 2:341–369.
Rogers, J.H. (1995), "Real shocks and real exchanges in really long-term data", Journal of International Economics 26, forthcoming.
Rogers, J.H. and M.A. Jenkins (1993), "Haircuts or hysteresis? Sources of movements in real exchange rates", mimeo, Pennsylvania State University.
Rogoff, K. (1992), "Traded goods consumption smoothing and the random walk behavior of the real exchange rate", Bank of Japan Monetary and Economic Studies 10:1–29.
Roll, R. (1979), "Violations of purchasing power parity and their implications for efficient international commodity markets", in: M. Sarnat and G. Szego, eds., International finance and trade (Ballinger, Cambridge, MA).
Samuelson, P.A. (1964), "Theoretical notes on trade problems", Review of Economics and Statistics 46:145–164.
Stockman, A. and L. Tesar (1995), "Tastes in a two-country model of the business cycle: Explaining international comovement", American Economic Review, 85:168–185.
Summers, R. and A. Heston (1991), "The Penn world table (Mark 5): An expanded set of international comparisons, 1950–88", Quarterly Journal of Economics 106:327–368.
Taylor, M.P. (1988), "An empirical examination of long-run purchasing power parity using cointegration techniques", Applied Economics 20:1369–1381.
Taylor, M. and P. McMahon (1988), "Long run purchasing power parity in the 1920s", European Economic Review 32:179–197.
Wei, S. and D. Parsley (1995), "Purchasing power disparity during the floating rate period: Exchange rate volatility, trade barriers and other culprits", mimeo, Harvard University.