NBER WORKING PAPER SERIES

INTERNATIONAL CAPITAL MOBILITY IN HISTORY: PURCHASING-POWER PARITY IN THE LONG RUN

Alan M. Taylor

Working Paper 5742

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1996

This work is related to a larger project with Maurice Obstfeld (Obstfeld and Taylor 1996) on the historical evolution of international capital mobility. I am indebted to Michael Bordo, Hugh Rockoff, and James Lothian for help with data enquiries, and to Peter Pedroni and Mark Watson for advice on econometric practice. For their helpful comments I wish to thank James Lothian, Michael Knetter, and participants at the NBER Conference on Exchange Rates (May 1996). Comments are very welcome. All errors are mine. This paper is part of NBER's project on International Capital Flows and NBER's research programs in the Development of the American Economy, and International Finance and Macroeconomics. It was presented at the NBER conference "Universities Research Conference on the Determination of Exchange Rates." We are grateful to the Center for International Political Economy for the support of this project. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper 5742
September 1996

INTERNATIONAL CAPITAL MOBILITY IN HISTORY: PURCHASING-POWER PARITY IN THE LONG RUN

ABSTRACT

This paper investigates purchasing-power parity (PPP) since the late nineteenth century for a sample of twenty countries, a broader sample of pooled annual data than has been studied before. Econometric results for time-series and panel samples allows us to test the robustness of the PPP hypothesis in different eras: the gold-standard, interwar, Bretton Woods, and the recent float. The evidence for PPP is mixed: Strong PPP, entailing stationarity of the real exchange rate, is not broadly supported, and real-exchange-rate dispersion shows counterintuitive historical patterns. However, not-much-weaker forms of PPP can be supported, with evidence of cointegration between different countries' common-currency price levels. Residual variances here confirm the conventional wisdom that the interwar period, particularly the Great Depression, represented the nadir of international capital market integration in the modern era.

Alan M. Taylor
Department of Economics
Northwestern University
2003 Sheridan Road
Evanston, IL 60208-2600
and NBER
amt@nwu.edu

1 PPP in Historical Context

This paper is concerned with the long-run evolution of international capital mobility in the nineteenth and twentieth centuries, and, as one dimension of this problem, re-examines the question of long-run purchasing-power parity (PPP). PPP can be identified as a necessary condition for the most stringent of capital market integration tests, the international equalization of real interest rates. Indeed, PPP in its Strong form, combined with speculative efficiency—uncovered interest parity (UIP) and covered interest parity (CIP)—is a sufficient condition for the equalization across countries of *ex ante* real interest rates.[1] Thus, PPP is indeed an important building block in any assessment of the historical success or failure of capital market integration measured by the (exacting) standard of real interest rate equalization.[2] And even PPP in weaker forms remains a valid benchmark and widely used criterion for judging the extent of successful international arbitrage, and a durable yardstick in theoretical and empirical analysis of the equilibrium real exchange rate (Isard 1995).

Of course, PPP remains a problematic concept for studies of capital mar-

[1]Formally, following Isard (1995), UIP entails that the forward-spot exchange rate differential equal the nominal interest rate difference between two currencies,

$$f_t - e_t = r_t - r_t^*, \tag{1}$$

where f_t is the log forward exchange rate, e_t is the log spot rate, and r_t and r_t^* are the home and foreign interest rates. CIP requires that the forward rate equal the expectation of the future spot rate,

$$E_t e_{t+1} = f_t, \tag{2}$$

where E_t denotes the expectation at time t. Finally, (Strong-Relative) PPP implies that expected exchange rate changes will equal expected inflation differentials,

$$E_t e_{t+1} - e_t = E_t \pi_t - E_t \pi_t^*, \tag{3}$$

where $\pi_t = p_{t+1} - p_t$ and $\pi_t^* = p_{t+1}^* - p_t^*$ are the inflation rates in the two currencies based on price levels p_t and p_t^*. It is straightforward to show that 1, 2, and 3 imply the equalization real *ex ante* real interest rates,

$$r_t - E_t \pi_t = r_t^* - E_t \pi_t^* \tag{4}$$

[2]For a historical study of real interest rate convergence in a cross-section of countries see Lothian (1995).

ket integration, since it may appear, at first sight, essentially a test of goods-market arbitrage. Though critics have worried about this problem, I will argue that the concern may be misplaced, especially in the longer run, both on theoretical and historical grounds. In theoretical terms, violations of PPP are often seen as stemming from barriers to goods market integration, whether due to transport costs or commercial policy. It is easy to see that in a simple two-good, general-equilibrium model, the predictions that trade barriers drive PPP deviations are not so clear cut. Small barriers to trade will reduce the trade volume incrementally (consider movements along offer curves) and drive wedges between home and foreign prices. Thus, home imports will be costlier than foreign exports—but foreign imports will be costlier than home exports too. Thus any aggregate price is a composite of import and export price indices, and will be subject to countervailing forces in both countries. In a symmetric, two-country world, with equal commodity weights in the price index, it is clear that overall valuations of PPP will be unaffected by barriers.[3] The distinction is between the general absolute price level and the structure of relative prices. Thus, thinking in terms of trade models, we are forced to confront goods-price determination in general equilibrium, a legitimate long-run concern. The same models also force us to confront factor price convergence (or equalization in the pure Heckscher-Ohlin model). Here again, the convergence of factor prices, such as the cost of capital, depends on the convergence of the *relative* price structure of traded commodities. These considerations urge us to reflect on aggregate PPP as an *absolute* price driven phenomenon. PPP is therefore somewhat distinct from the pure law-of-one-price concept applied to commodities, and, I argue, more properly in the domain of monetary economics and the macroeconomic theory of inflation.

On historical grounds, we would rapidly arrive the same conclusion. Conceiving of PPP in 1922, Cassel was motivated by the vast dispersion in national price levels driven by wartime inflations in the various belligerent countries. Adjusting exchange rates to be consistent with PPP, so as to reinstate the gold standard credibly, was quickly seen as a macroeco-

[3]Explicitly, let the *ad valorem* cost of the barrier be t, and symmetric on trade flowing in both directions between two identical countries, A and B, which each produce two goods, X and Y. Suppose A (B) exports X (Y). Then, $P_{BX} = P_{AX}(1+t)$ and $P_{AY} = P_{BY}(1+t)$. Let the goods have equal weights in price indices $P_i = P_X^{\alpha_X} P_Y^{\alpha_Y}$ for countries $i = A, B$. Then, by symmetry, it is clear that $P_A = P_B$. Thus, for any barrier t, the absolute price levels of the countries are equal.

nomic problem requiring monetary stringency. In the interim, convertibility and disequilibrium exchange rates were sustainable only with strict controls which appeared in the war, subsided through the 1920s, but rose to unprecedented heights in the 1930s and persisted into the 1960s, 1970s, and 1980s under the Bretton Woods system. And although overvaluations, say, might be contained with indirect controls to stem capital outflows, such as import restrictions, inevitably such controls had a sting in the tail: Lerner symmetry promised a countervailing contraction of exports which would exacerbate the initial balance-of-payments problem (Einzig 1934; Nurkse 1944; Yeager 1976). Thus, exchange rate equilibration under gold standard "rules of the game" was potentially feasible, but often blocked by political fear of the costs of adjustment in countries which needed to deflate, or by lingering nightmares of hyperinflation in countries which needed to reflate (Eichengreen 1992). In such circumstances, absent a willingness to devalue, the sustainability of deviant fixed exchange rate parities—albeit only on the official market—required an ever-more convoluted system of controls: the direct supervision of exchange transactions, the rationing of foreign exchange by transactions, the creation of "blocked balances" to manage trade balances along bilateral lines. The legacy of this implosion of the international monetary system was a cautious approach under Bretton Woods to the "destabilizing" flows of "hot" money which had wreaked havoc in the inter-war period, and the understandably slow removal of capital controls after 1945 (Obstfeld and Taylor 1996). Viewed in such terms, the historical success or failure of PPP can be seen as intimately tied to the mobility of global financial capital in the course of the twentieth century, and such is the starting point for the present paper.

What new findings can this paper claim to offer? From a historical standpoint, there have been numerous studies of PPP for various countries over the period in question, some covering a particular era or monetary regime. McCloskey and Zecher (1984) argued that PPP worked very well under the Anglo-American gold standard before 1914, using this as a basis for a monetary theory of gold-standard adjustment. Diebold, Husted, and Rush (1991) explored a very long run of nineteenth century data for six countries, and found support for PPP based on the low-frequency information lacking in short-sample studies. Abauf and Jorion (1990) studied a century of dollar-franc-sterling exchange rate data and verified PPP; Lothian and Taylor (1996) found the same for *two* centuries of dollar-franc-sterling data. Lothian (1990) also found evidence that real exchange rates were stationary for Japan, the U.S., the U.K., and France for the period 1875-1986, although yen exchange rates exhibited only trend-stationarity—an oft-repeated finding that the real yen exchange rate has appreciated over the long run against *all* currencies. In full length monographs, both Lee (1978) and Officer (1982) found strong evidence in favor of PPP based on analysis of long time-series running from the pre-1914 gold standard to the managed float of the 1970s. Obviously, this paper builds on a very strong foundation of historical work by a number of scholars, covering various countries in different time periods.[4] Here, I aim to consolidate and expand these approaches in a unified framework, and using the latest techniques, to examine the applicability of PPP from circa 1880 to the present in a broad cross-section of countries.

Of late, new techniques have appeared in abundance. The idea of PPP dates back to Cassel's (1922) argument that, save for short run deviations, well-functioning and integrated international markets should produce a tendency for the prices of baskets of goods denominated in a common currency to converge–a seemingly mundane extension of the law of one price across countries. However, in their recent comprehensive review of the purchasing-power parity literature, Froot and Rogoff (1995) could declare that what was a "fairly dull research topic" only a decade ago has recently been the focus of substantial controversy and the subject of a growing body of literature.

Recent empirical research, mostly based on the time-series analysis of short spans of data for the floating-rate (post-Bretton-Woods) era had led many to conclude that PPP failed to hold, and that the real exchange rate followed a random walk, with no mean-reversion property. However, more recent studies have challenged this conventional wisdom, seeing it as a flawed result arising from lack of statistical power, a consequence not only of the small-sample short-run series employed, but also of the inherent weaknesses of standard unit-root tests. A newly emerging literature exploits more data and higher-powered techniques, and claims that, in the long run, PPP does indeed hold: it appears from these studies that real exchange rates exhibit mean reversion with a half-life of five years or so (M. P. Taylor 1995; Froot and Rogoff 1995).

The newer findings use various steps to expand the size of samples used to test PPP. As noted, it has been possible to use much longer-run time series for certain individual countries, spanning a century or more; typically such exercises have concentrated on more-developed countries with good historical

[4]Other studies of long run data are numerous (Frankel 1986; Edison 1987; Johnson 1990; Glen 1992; Kim 1990).

data availability (e.g., U.S., Britain, France). Alternatively, researchers have expanded the data for the recent float or postwar periods cross-sectionally to exploit the additional information in panel data (Wei and Parsley 1995; Frankel and Rose 1995; Pedroni 1995).

It is still too early to say whether the revisionist PPP findings will prove robust, and already challenges to this interpretation have emerged. One may find fault with the ways in which cross-section information and panel methodologies have been applied (O'Connell 1996). Some have noted that the inferences based on panel methods are sensitive to sample selection (Papell 1995). In particular, results appear sensitive to the choice of base country, e.g. the U.S. versus Germany (Papell 1995; Wei and Parsley 1995; Edison, Gagnon, and Melick 1995). Others caution that detecting a unit root in time series may be complicated by the fact that price indices can be viewed as the sum of a stationary tradable relative-price component and a non-stationary non-tradable relative-price component (Engel 1996)—a time-series finding which echoes the venerable Balassa-Samuelson objection to the pure PPP hypothesis based on differential rates of productivity growth in traded and non-traded goods sectors (Balassa 1964; Samuelson 1964). The distinction of the present study is to introduce the recent wave of empirical innovations to a longer span of historical data, both to investigate the robustness of the recent findings and to explore the historical evolution of PPP.

Can the question of long-run PPP helpfully inform, and be informed by, a new historical study such as this? I believe so. First, history has a lot to say about the robustness of the PPP hypothesis. We can use historical evidence to expand the power of tests by enlarging both the sample size and the variation in the data. To that end, I have constructed a panel data set covering annual time series for twenty countries since 1880, a sample unlike any other in length and breadth. I use this data to evaluate PPP using traditional approaches and the latest panel econometric techniques. By focusing on various subsamples—like the classical gold standard, the interwar period, Bretton Woods, and the recent float—I can further explore the validity of the PPP hypothesis under different monetary regimes.

Second, as already noted, the PPP hypothesis has something to say about history, specifically our understanding of the evolution of global capital market integration. The textbook position on the fluctuations in international capital mobility over the last century is well known: a presumption of a high degree of integration before 1914, disintegration through the two world wars and the Great Depression, and a gradual re-integration of the system under Bretton Woods, with further integration (despite increased volatility) under the recent float. This characterization of history, I argue, ought to be manifested in empirical measures of market efficiency and integration, and this PPP study forms one piece of evidence to either support or refute the textbook view. In particular, we want to know whether empirical evidence on PPP justifies the common view that the Great Depression represented the "low-water mark" of international capital mobility since the mid-nineteenth century.

I proceed as follows. Section 2 briefly surveys the relevant theoretcial and empirical literature. Section 3 attempts to find evidence in my long-run sample of "Strong" PPP , a relationship where, for each country, world and domestic prices have a proportional relationship. The evidence here is unfavorable. Section 4 relaxes the restriction and tests for so-called "Weak" PPP. The evidence here is more favorable, and an analysis of deviations from PPP thus measured confirms the textbook characterization of phases of integration and disintegration in the world capital market. Section 5 concludes.

2 Methodological Perspectives on PPP

A full review of the methodological issues surrounding the investigation of PPP is not warranted here, and the reader may consult, for example, the recent comprehensive study by Froot and Rogoff (1995) on which I draw.[5] However, a basic framework will be established in this section, with terminology laid down, and some complications considered.

A starting point is the notion of *Absolute Consumption-Based PPP*, which may be written several different ways,

$$\begin{aligned} p_t &= p_t^* + e_t; & (5) \\ e_t &= p_t - p_t^*; & (6) \\ p_t - e_t &= p_t^*; & (7) \end{aligned}$$

where p_t is the log domestic-currency consumption-price level, p_t^* is the log foreign-currency consumption-price level, and e_t is the log of the exchange rate (domestic-currency price of foreign exchange).[6]

[5] Further survey material can be found in M. P. Taylor (1995) and Isard (1995).

[6] Equation 5 says that domestic-currency prices should equal foreign-currency prices converted at the exchange rate. Equation 6 says that the exchange rate should reflect dif-

However, it is rare to have absolute national price-level data available.[7] More commonly, PPP is studied using national price-level indices measured in arbitrary, non-commensurate units. This usually enforces a retreat to the testing of *Relative Consumption-Based PPP* in various analogous forms

$$\Delta p_t = \Delta p_t^* + \Delta e_t; \tag{8}$$
$$\Delta e_t = \Delta p_t - \Delta p_t^*; \tag{9}$$
$$\Delta p_t - \Delta e_t = \Delta p_t^*; \tag{10}$$

where now only changes in relative price levels are linked to the change in the exchange rate. An alternative to these formulations is to test equations like 5 with the inclusion of a constant term,

$$p_t = C + p_t^* + e_t; \tag{11}$$
$$e_t = -C + p_t - p_t^*; \tag{12}$$
$$p_t - e_t = C + p_t^*. \tag{13}$$

Equations of this form, such as equation 12, invite time-series testing using a simple estimating equation of the following form (e.g., Frenkel 1981),

$$e_t = \alpha + \beta(p_t - p_t^*) + \epsilon_t. \tag{14}$$

A traditional early test of Relative PPP using 14 amounted to a test of the restriction $\beta = 1$. Several pitfalls are apparent with this methodology. First, the residuals ϵ_t must be assumed to be stationary for standard inference to be valid. Second, the variables e_t and p_t (if not also p_t^*) might reasonably be considered endogenous, and simultaneously determined. There is no *a priori* reason to have exchange rates on the left-hand side and prices on the right. Lastly, there are reasons to expect that, even if markets are well-integrated, it may not be the case that $\beta = 1$. The Balassa-Samuelson hypothesis, as well as other theories, note that countries cannot buy each other's consumption basket. Rather, non-traded goods circumvent the exact pass-through of price shocks from one market to another. This might plausibly lead to a long-run relationship of the form 14, but with a slope not equal to unity. The term

ferences in countries' national price levels. Equation 7 says that prices should be equalized when denominated in a common (here, the foreign) currency.

[7] An exception is the ICP data of Heston, Summers, et al. (1994).

Weak PPP may be used to describe the case $\beta \neq 1$, as opposed to *Strong PPP*, the case $\beta = 1$.

An alternative methodology *imposes* the restriction that $\beta = 1$ (Strong PPP), in which case the residuals can be constructed, rather than estimated, using equation 14. The residuals are then termed the *real exchange rate*, the log of which is written $q_t = e_t - p_t + p_t^*$. In this methodology attention focuses on whether this variable is stationary, or else obeys the null of a random walk. Such tests often amount to a test for mean reversion in the equation

$$\Delta q_t = \alpha + \beta q_{t-1} + \epsilon_t, \tag{15}$$

where $H_1 : \beta < 0$ is evidence of stationarity. Failure to reject the null $H_0 : \beta = 0$ using the standard battery of unit root tests is viewed as a failure of the PPP hypothesis because q_t then exhibits no tendency toward mean reversion. However, the weak power of unit-root tests gave rise to doubts, despite the almost universal acceptance of the unit-root null in contemporary data. Even if the alternative H_1 were true, with slow mean reversion it could take over a century of data to reject the unit root null with conventional tests (Frankel 1986; 1990). Further, it seemed unwarranted to impose the restriction $\beta = 1$ which, in effect, prespecified the cointegrating vector for (e_t, p_t, p_t^*). Taken together these criticisms seemed to beg for the introduction of a cointegration approach, and a search for increased sample size. Such has been the thrust of recent contributions to the literature .

In the cointegration approach an equation such as 14 could be estimated, the residuals could be examined for stationarity, the cointegrating vector (α, β) could be inferred, and the restrictions implied by Strong versus Weak PPP could be tested rather than assumed. In fact, even looser specifications of the form

$$e_t = \alpha + \beta_1 p_t - \beta_2 p_t^* + \epsilon_t \tag{16}$$

could be estimated, with no prior restriction on $(\alpha, \beta_1, \beta_2)$. Again, objections such as Balassa-Samuelson, or problems of mismeasurement (bias) of price indices, could lead to situations where $\beta_1 \neq 1$ or $\beta_2 \neq 1$. However, in this respect the cointegration approach's flexible specification is also its downfall, as estimation of such equations on the recent float produce unlikely cointegrating vectors, with β_1 and β_2 ranging wildly. Froot and Rogoff (1995) ascribe this to poor-fit cointegration biases. But Kim (1990) shows that these problems are diminished when historical data are used. I will later arrive at a similar finding with broader historical panel data.

One way to increase the power of the tests is to expand the sample size, and the use of panel data was one obvious way to accomplish this aim. However, this gives rise to further econometric problems, and the techniques for the analysis of unit roots and cointegration in panel data are still in their infancy. Frankel and Rose (1996) attempted to form a panel version of equation 15, using postwar data (during and after Bretton Woods) for a very wide panel of more than 100 countries They construct the real exchange rate so as to implicitly test the Strong-Relative variant of PPP, estimating an equation of the form

$$\Delta q_{it} = \alpha_i + \beta q_{i,t-1} + \epsilon_{it}. \tag{17}$$

Here, the critical values of β are *not* the same as those in the univariate time-series case. The appropriate critical values for the "t-like" tests were derived by Levin and Lin (1992). Unfortunately, here again the unit-root null cannot be rejected.[8] Indeed, some now question the newly emerging view that PPP might hold after all. Papell (1995) has found that acceptance of mean reversion is very sensitive to the size of the panel constructed. O'Connell (1996) notes various problems, including the benchmark issue previously noted.

In theory the choice of benchmark country (for p^*) should be immaterial, but this matters in practice (e.g., it is easier to reject the unit-root null during the recent float when Germany is the benchmark rather than the U.S., a finding possibly related to the unusual swings in the dollar in the 1980s). A solution to this problem in the panel context is, instead, to set the benchmark relative to a world-average basket of currencies. Thus O'Connell (1996) suggests estimating a variant of equation 13 with the "dollarized" price index as the variable to be studied in country i, defined as $r_{it} = p_{it} - e_{it}$, which is supposed to be following a mean-reverting process with respect to an unobserved "world nominal anchor" a_t, so that

$$\Delta(r_{it} - a_t) = \rho(r_{i,t-1} - a_{t-1}) + \epsilon_{it}. \tag{18}$$

We can set the benchmark relative to the U.S. by subtracting equation 18 for the U.S. from equation 18 for country i, eliminating a_t to obtain

$$\Delta(r_{it} - r_{US,t}) = \rho(r_{i,t-1} - r_{US,t-1}) + \epsilon_{it} - \epsilon_{US,t}. \tag{19}$$

The benchmark-related problems are clear. First, the term $\epsilon_{US,t}$ introduces cross-sectional dependence in the panel estimation. Second, if ρ is not constant across countries then the benchmarking introduces serial correlation in

[8]See O'Connell (1996).

the error terms of 19. O'Connell (1996) therefore advocates subtracting a world average of 18 from itself, implying

$$\Delta(r_{it} - r_{Wt}) = \rho(r_{i,t-1} - r_{W,t-1}) + \epsilon_{it} - \epsilon_{Wt}, \tag{20}$$

where a subscript W denotes a world average.[9]

Finally, we should again recall that the Strong PPP tests are really joint hypothesis tests. Moreover, in a panel setting, the imposition of restrictions on coefficients across many members of the panel might further encourage rejection of a cointegrating relationship. Specifically, all the above panel estimators assume that the panel is homogeneous in its slope parameters (though not in its intercepts). This presupposes Strong PPP. If, instead, the slope coefficients vary across countries, then a more flexible heterogeneous panel estimator is called for, and appropriate tests for cointegration need to be devised for this context. The residual has to be estimated, not constructed, and one can no longer speak of q_{it} reverting to a target. Pedroni (1995) has developed just such an approach, and has found the appropriate cointegration tests needed for the heterogeneous case, illustrating them with an application to PPP in the recent float. His estimating equation is

$$e_{it} = \alpha_i + \beta_i(p_{it} - p^*_{it}) + \epsilon_{it}, \tag{21}$$

where the benchmark country is the U.S. (Note that a symmetry restriction is imposed, in that domestic and foreign prices enter with the same coefficient.) An alternative heterogeneous panel-estimating equation following O'Connell's (1996) "world benchmark" method would take the form

$$r_{it} = \alpha_i + \beta_i r_{Wt} + \epsilon_{it}. \tag{22}$$

The latter approach will be important in our study, and its appeal is twofold. First, it does not utilize a single-country benchmark, which is important if certain countries' trade baskets are unusual or if the benchmark country has peculiar in-sample exchange rate movements, as suspected. Second, it allows for full heterogeneity in the individual-country dynamics, an added flexibility in form which may give greater power to reject the no-cointegration null. To

[9]Here we are testing the relationship between country i's "dollarized" price index and that of the world. This may make it easier to maintain an assumption of cross-sectional independence since the term $\epsilon_{W,t}$ only has an impact on the disturbance covariance matrix that is $O(1/\sqrt{N})$ for a panel of cross-sectional width N.

see why this is important for the present study we will first need to examine the shortcomings of the more restrictive Strong PPP specifications when applied to long-run historical experience.

3 Data

A brief word is in order about the database used in this study. It consists of time-series estimates of annual average exchange rates (relative to the U.S. dollar) and price levels (consumer price deflators, or, when they are not available, GDP deflators) for twenty countries covering the period 1880 to 1994. For its construction I have relied on standard sources. The principal published sources are the statistical volumes of Brian Mitchell (*International Historical Statistics*). For the convenient provision of previously electronically-compiled data from these and other sources I am grateful to Michael Bordo. Essentially, all such series build on the individual labors of national statistical offices and individual scholars dedicated to the compilation and revision of national macroeconomic statistics. Like so many, I am indebted to this devoted group of researchers. A full documentation of these sources is provided in a data appendix at the end of this paper, and the complete set of data is available from the author upon request.

4 Weaker Evidence for Strong PPP

I begin with the assumption of Strong PPP, namely that the log real exchange rate $q_t = e_t - p_t + p_t^*$ has a long-run tendency to revert toward some equilibrium value. In any sample this should be roughly the sample mean of the data. We can test for this first using standard unit-root tests on individual country time series. For additional power we can employ panel methods. In fact, it proves difficult to reject the null that q_t has a unit root. A direct examination of the long-run trends in the levels and dispersion of q_t for our sample of twenty countries since 1880 suggests that this conclusion is unsurprising.

4.1 Time-series tests of a unit root in the real exchange rate

I first perform tests for mean-reversion in the real exchange rate on a country-by-country basis. Table 1 presents a summary of unit root and cointegration tests applied to the log real exchange rate q_t. To maximize the power of the tests I apply them to the longest timespan of data available for each country. I apply the test to two variants of the log real exchange rate, one defined with U.S. benchmarking, one with a "rest-of-world benchmark"

$$\begin{aligned} q_{it} &= -r_{it} + r_{US,t}; && (23) \\ q_{it}^{ROW} &= -r_{it} + r_{ROW,t}; && (24) \end{aligned}$$

where q_{it}^{ROW} is calculated for a maximum-length 16-country rectangular sub-panel (this benchmark could not be applied to all 20 countries because of some missing data for certain countries before 1914). Here, *ROW* refers to an average taken over all countries other than i.

The unit root tests reveal two important problems in the study of long-run PPP. First, if strong PPP is the reference point, it is unlikely to be satisfied for all countries at all times. Only a few of our sample countries exhibit stationarity in the log real exchange rate as tested using the standard Dickey-Fuller (DF) tests.[10] Second, the benchmark choice matters: the results are sensitive to whether the U.S. or the world is used as the reference for foreign price levels. Based on either benchmark, however, only a handful of countries satisfy the stationarity assumption.

Since standard unit root tests have been criticised in applications such as this, I also tried potentially more-powerful cointegration tests: the error-correction model (ECM) test proposed by Kremers, Ericsson, and Dolado (1992), and the Horvath-Watson (1995) test.[11] In the ECM test, a restricted ECM regression is formed with the cointegrating vector prespecified such that the error-correction term is just the real exchange rate q_t, namely

$$\Delta r_{it} = \alpha + \beta q_{i,t-1} + \gamma \Delta r_{*,t} + \epsilon_{it}. \quad (25)$$

In the present, I restrict attention to the single-lag case. The test depends on the t-statistic of the error-correction coefficient β. It can be seen that some

[10] Philips-Perron tests gave similar results and are omitted.

[11] Both such tests were used by Engel (1996). Like him, I take critical values of the ECM test from Hansen (1995), estimating a nuisance parameter.

power is added to the test in Table 1, and more series appear cointegrated with the ECM test. An even more general specification was attempted with the Horvath-Watson test, allowing the 3×1 vector $y_t = (p_t, e_t, p_t^*)'$ to be cointegrated according to a prespecified vector $(1, -1, -1)$, but otherwise subject to a general vector autoregression process. Here I work only with the U.S. benchmark data. I estimate a vector error-correction model (VECM) of the form

$$\Delta y_{it} = \alpha + \beta q_{i,t-1} + \gamma_1 \Delta y_{*,t-1} + \gamma_2 \Delta y_{*,t-2} + \cdots + \gamma_p \Delta y_{*,t-p} + \epsilon_{it}, \tag{26}$$

were, α, β, and ϵ_{it} are 3×1; γ_j is 3×3; and the lag length p is chosen by a step-down procedure.[12] The test depends on the significance of the error-correction coefficient vector β, evaluated by a Wald test. Looking at Table 1, this test proves a little more powerful, but still 6 out of 19 countries fail the test. Moreover, viewing all the tests of Strong PPP, there seems to be no consistent pattern at the country level.

4.2 Panel tests of mean reversion in the real exchange rate

An optimistic interpretation could be that Strong PPP is sometimes verifiable, but not universally so. Yet based on a sceptical view of the results in Table 1, a presumption of the general validity of Strong PPP in the long run probably cannot be justified. This may not be as damaging as at first appears. Perhaps this failure merely results from my attempt to use "only" one hundred years of annuyal time series data. For as we have seen, conventional test may prove weak on such samples. Can more powerful panel techniques overturn this negative conclusion?

Table 2 presents tests based on the panel methodology of Frankel and Rose (1995), whereby I estimate equation 17 and seek a significant and negative coefficient β as evidence of mean reversion, and, hence, stationarity in q_{it}. Implicitly we are estimating a restricted panel equation, with homogeneous slopes (β is the same for all countries i), but with heterogeneous intercepts—since in the Frankel and Rose (1995) methodology each series q_{it} has the country mean removed. The critical values are provided by Levin and Lin (1992).

[12]This again follows Engel (1996).

The results are disappointing and offer little support for the Strong PPP hypothesis over the long run. To counteract concerns that these weak results may follow from an unjustified pooling of the data across various regimes, I also apply the test for various subsamples: the pre-1914 gold standard (G), the interwar period 1914-1945 (I), the Bretton-Woods era 1946-1971 (B), and the recent float post-1971 (F). It is noticeable that splitting the sample reveals much larger speeds of convergence (0.11 to 0.25 per annum) and raises the significance level of the test statistics. This warns us that pooling across the sample might be unwarranted. In one case, mean reversion is supported, namely under the gold standard using world benchmark data, and *a priori* we might have suspected this to be the easiest case to verify. There is also the marginal case of the recent float. The latter is an odd case to have: the float is focused on by Frankel and Rose, and many other recent contributors in the "pro-PPP" revisionist school. It would be contrary to the received wisdom, then, if the float were the sample in which PPP were *relatively easy* to uphold, as most research has concluded that the PPP hypothesis ought to be most easily validated under fixed-exchange rates, that is, in less volatile monetary regimes *other than* the recent float (Froot and Rogoff 1995).

So, again, based on a presumption that we should be testing for Strong PPP, we cannot reject the hypothesis that there is no long-run equilibrium level of the real exchange rate based on historical experience since the late nineteenth century. Why should this problematic finding appear?

4.3 Historical patterns of deviations from Strong PPP

Figures 1 and 2, and Table 3, suggest some reasons why the restrictive requirements of Strong PPP might be an unreasonable standard for our tests. Figure 1 shows a measure of the dispersion of the log real exchange rate $\sigma_q^2(t)$, defined as the variance in the log de-meaned real exchange rate across all countries i at a given time t,

$$\sigma_q^2(t) = \mathrm{Var}(q_{is} - q_{iT} | s = t) \tag{27}$$

where q_{iT} is the mean of q_{it} over all t. Thus, $\sigma_q^2(t)$ is a measure of the extent to which Strong PPP is violated in our sample at a given point in time t. We would expect this to reflect phases of integration and disintegration on world markets. When markets are perfectly integrated, the presumption is that PPP will hold with perfect equality. In that case $q_{it} = q_{iT}$ for all i, t, and dispersion is zero. Any deviation from $q_{it} = q_{iT}$ will be rapidly eliminated by arbitrage. Conversely, if markets are poorly integrated, there is reason to expect deviations from PPP to arise and show persistence.

Figure 1 and Table 3 indicate that deviations from long-run mean real exchange rates have been most dramatic in the contemporary era, a finding which clashes with conventional interpretations of long-run market integration. We see no evidence of large real exchange rate dispersion around the two wars or during the dislocations of the interwar period. These are odd findings if accepted at face value. Is it time to tear up the textbook historical accounts of the evolution of international capital mobility? Was the interwar period truly an era of historically *low* deviations from parity? Were Cassel and Keynes, or Poincaré, somehow mistaken in their perception of the severity of parity dislocations after World War One? And did the many and varied devaluations following the collapse of the gold standard in the 1930s really have so little impact of the dispersion of real exchange rates? I would say not. I will argue that the above evidence is of dubious value, for the simple reason that tests of Strong PPP over the long run seem to fail. In such cases the use of mean real exchange rates as if they were a target of mean reversion, and, hence, a relevant benchmark for dispersion calculations, seems unwarranted. As we will see below, when deviations are measured relative to a Weak PPP benchmark, the long-run evidence on deviations from PPP appears more credible.

Figure 2 helps us diagnose how these strange results arose. The charts plot q_{it} (the black line) and q_{it}^{ROW} (the gray line) for all countries since 1880. It is apparent that the real exchange rate is far from stationary in several of our sample countries over a hundred-year timespan. Ocular inference suggests that Strong PPP is likely to be violated for Japan, Switzerland, Brazil, and Portugal.[13] In this mostly OECD sample, the first two countries are noted for high productivity, the latter two for low productivity performance, suggesting possible Balassa-Samuelson effects. The inclusion of countries with marked long-run real exchange rate trends almost inevitably leans to mid-sample minima on measures of cross-sectional dispersion about the long-

[13]The long-run drift of the Japanese real exchange rate is well know (Lothian 1990). The presence in our database of countries such as Argentina, Chile, Brazil, and Portugal guards against the "survivorship" bias in studies which focus only on exchange rates in more-developed countries. The more-developed countries may provide stronger support for PPP for structural reasons related to stable policies, market institutions, and other structural differences (Froot and Rogoff 1995).

run mean. This explains the curious pattern of dispersion in Figure 2. The mid-sample period is the interwar period, so such findings are forced to jar with the conventional wisdom. The result follows from adherence to Strong PPP, and the prespecification of what the long-run equilibrium target is for each country. This approach sidesteps the problems of nontradables and productivity, not to mention the incommensurability of international price indices—issues which might easily produce long-run drift on theoretical or empirical grounds. Thus confining attention to the Strong PPP case is unlikely to produce positive support in all cases studied, and this implicit restriction could, as part of a joint hypothesis test, be the reason for a failure to reject the unit root null.

5 Stronger Evidence for Weak PPP

In this section I retreat to the Weak PPP specification since the above tests so often reject Strong PPP in very long-run data. I begin with some univariate time-series tests based on an error-correction model (ECM) which offer some support for the cointegration property. I then exploit heterogeneous panel methods to increase the power of the cointegration test. I find strong evidence to support Weak PPP, with a good indication that the coefficients of domestic and foreign prices do not enter the cointegrating vector symmetrically.

5.1 Time-series ECM tests of Weak PPP

For country i let us presume that equation 22 holds such that domestic and foreign dollarized price indices are cointegrated according to an equation $r_{it} = \alpha_i + \beta_i r_{*t} + \epsilon_{it}$ where r_{it} and r_{*t} are $I(1)$ and ϵ_{it} is $I(0)$, and where $*$ may represent US with benchmarking relative to the U.S. or ROW with benchmarking relative to the rest of the countries in the sample.

If such a cointegrating relationship holds then the dynamic relationship can also be expressed in error-correction form (Engle and Granger 1987). The simplest such form is the first-order error-correction model (ECM), and a feasible test for cointegration here is the "t-like" test based on the error-correction term (Kremers, Ericsson, and Dolado 1992). In this situation we may posit an ECM relationship, form

$$\Delta q_{it} = \alpha + \beta q_{i,t-1} + \gamma r_{*,t-1} + \eta \Delta r_{*,t} + \epsilon_{it}. \qquad (28)$$

We need to assume that r_{*t} is weakly exogenous for country i, which may not be unreasonable (at least for a small open economy). It is then the case that this relatively simple form for the structural dynamics allows much greater flexibility than a pure mean-reversion or cointegration test, whilst still allowing us to test for both Weak and Strong PPP. Cointegration holds when β is significantly less than zero. It can be shown that the OLSt-statistic of β (t_{ECM}) is asymptotically distributed as a normal $N(0,1)$ random variable. Moreover, the power of this test is superior to the standard univariate tests (e.g., Dickey-Fuller) because information in the covariate term $\Delta r_{*,t}$ is not discarded. If cointegration holds then either Weak or Strong PPP may hold. In the special case where $\gamma = 0$ Strong PPP holds and the long run cointegrating vector for (r_{it}, r_{*t}) is $(1, -1)$. If $\gamma \neq 0$ then Weak PPP holds and the cointegrating vector is not symmetric in prices. The coefficient η introduces richer short-run dynamics and has a natural interpretation. If $\eta = 0$ then short-run pass-through from foreign price shocks to domestic price shocks is one-for-one. If $\eta > 0$ then there is overshooting, and if $\eta < 0$ there is undershooting.

Table 4 shows estimates of these equations. Several implications follow. First, the global restriction that the cointegrating vector be $(1, -1)$, implicit in Strong PPP analysis, appears to be unjustified, as γ varies considerably and is often nonzero. Second, the more flexible form captures interesting short-run dynamics: in the several cases studied we can examine η and find evidence of both undershooting and overshooting. Third, and very importantly, the cointegration tests based on t_{ECM} provide much stronger support for a long-run equilibrium relationship between r_{it} and r_{*t}. Fourth, from an examination of the β estimates we see that the speed of convergence to equilibrium varies considerably from country to country, further illustrating the extent of cross-country heterogeneity in the panel. In the cases where β is significant, however, the half life of deviations is usually less than five years ($\beta < -0.13$), in keeping with recent findings.

Lastly, it is worth noting that the cointegration finding is consistent with the general result that weakening the proportionality and symmetry restrictions makes it easier to find stationary residuals (Froot and Rogoff 1995). However, Froot and Rogoff also caution that this finding usually comes at a heavy cost: in the literature, the unrestricted cointegrating vector usually deviates markedly from symmetry and proportionality, defying easy interpretation and therefore casting doubt on the empirical framework. Here, we have not had to pay so heavy a price. The size of the coefficient γ measures the deviation from proportionality between home and domestic prices in the annual time series. It can be seen that γ mostly lies in the range [-0.05,0.05] meaning a proportionality coefficient between prices in the range [0.95,1.05].[14] Such modest deviations seem entirely reasonable given the various definitions of price indices across countries, the certain non-homogeneity of price indices across time in any one country, and the plausible supposition that non-tradable components may suffer from the Balassa-Samuelson effects already noted. Thus, although Strong PPP is ultimately rejected in favor of Weak PPP in this study, the distinction is not so large.

5.2 Panel cointegration tests of Weak PPP

If cross-country heterogeneity must be admitted to the analysis of panel dynamics, then attempts to raise the power of cointegration tests by using a panel instead of individual time series must take into account the modifications necessary to construct valid test statistics and finite sample distributions for such special cases. The Levin and Lin (1992) study only provides cointegration tests for panels with homogeneous slopes. Here, in the case of heterogeneous slopes I appeal to the work of Pedroni (1995), and estimate an equation of the form 22 where $r_{it} = \alpha_i + \beta_i r_{ROW,t} + \epsilon_{it}$.

Table 5 reports the results for our balanced panel of sixteen countries using the "rest-of-world" benchmark for the foreign price index. With the increased sample size we are able to perform this test not only for the entire sample period, but also for the four subperiods (G, I, B, F) and sequential groupings thereof (GI, IB, etc.). The results are favorable to the hypothesis of Weak PPP. They are consistent with Pedroni's (1995) findings for the recent float, and extend his results to earlier epochs. Only in the interwar period do we fail to reject the null of no cointegration, and this may be due to small sample size (few observations in the time dimension). In every other case, including periods which overlap the interwar period, the null is rejected. Table 6, finally, shows estimates of the heterogeneous panel equation 22 for the full sample in the nineteenth and twentieth centuries. Intercept terms have no interpretation, and are omitted for clarity. The slope terms are indeed heterogeneous. The slopes which differ most from unity are the four cases we identified from the long-run trends in real exchange rates: Japan,

[14]There are only two exceptions: Chile (-0.11) using the U.S. benchmark, and Brazil (-0.07) using the ROW benchmark.

Switzerland, Brazil, and Portugal.

Clearly, by this reckoning, we again find that world price indices do not feed one-for-one into domestic price indices. But again, as in the previous section, we should note that deviations from Strong PPP are not marked by the standards of the literature, as the proportionality of domestic and world prices falls in the range [0.88,1.20] for all but four countries.[15] For the same reasons as above, deviation from pure symmetry (Strong PPP) are anticipated here too. The deviations are stronger than those measured using the univariate ECM framework, possibly as a result of the additional restrictions on the error structure imposed by the panel methodology, a price we pay for the additional power of this method to identify cointegration. The ECM framework also differs in focusing on the adjustment process for changes in r rather than an equilibrium relationship in the levels of r. Understandably, the results differ in the two approaches.

5.3 Historical patterns of deviations from Weak PPP

Thus, we favor Weak PPP, but find the relationship not qualitatively so different from Strong PPP: the above estimates of the slope coefficients, the differences $(\beta_i - 1)$ are not unusually large in most cases. Historical interpretation has never turned on such criteria. Indeed, McCloskey and Zecher's (1984) argue that significant differences from unity are not the true criterion for judging PPP: significance only reflects defeat of the sampling problem. We can still accept these results as evidence of long-run arbitrage and market integration maintaining an equilibrium link between prices in spatially separate markets: it may not matter quantitatively or qualitatively whether (statistically) we can say that β_i equals 0.9 or 0.99, or 0.999999. Indeed, the confidence interval around β_i may not be the prime focus of historical interest.

Rather, we might profitably ask, given the cointegrating regression as a best-fit estimate of the (Weak PPP) equilibrium relationship, how far from this relationship did the international system deviate at certain times? That is, what do the residuals of the regression tell us about the size and persistence of PPP deviations? This evidence on the adjustment process embodies both concerns about the speed of reversion to equilibrium (propagation) and about the size of shocks to the system (impulses). And it recognizes the need to

[15] These four are Switzerland (1.64), Japan (1.50), Brazil (0.49), and Portugal (0.65).

not only examine the estimated speeds of adjustment (cf. β convergence in the language of growth models) but also the dispersion of the deviations themselves (cf. σ convergence). I now turn to examine the dispersion of the panel regression residuals as a measure of diseqilibrium: can they tell a more reasonable story than the dispersion of the real exchange rate?

Table 7 and Figure 3 offer a basis for evaluating deviations from Weak PPP. The residuals ϵ_{it} estimated in equation 22 now provide a measure of the deviation of log "dollarized" price levels from their equilibrium value. Hence, we may usefully study the dispersion of these residuals over time to get some sense of the extent to which deviations from Weak PPP equilibrium waxed and waned in different eras. Figure 3 shows a measure of the dispersion of the residuals $\sigma_\epsilon^2(t)$, defined as

$$\sigma_\epsilon^2(t) = \mathrm{Var}(\epsilon_{is}|s = t). \tag{29}$$

This variance is presented in two forms, one of which includes in the regression additional interecpt terms for each subperiod (G,I,B,F), to correct for the possible temporal inconsistency of the price series—although I will focus on the upper line in the chart, the uncorrected measure. [16]

Unlike the earlier measure of real exchange rate dispersion, $\sigma_q^2(t)$ shown in Figure 3, this measure of deviations from parity produces record of international market integration that is at once credible, yet useful in its quantitative foundation and capable of providing new insights. PPP deviations were extremely small in the pre-1914 era, as is well appreciated from our knowledge of the gold standard. Dispersion grew dramatically after 1914 with the collapse of the classical gold standard and the disparate inflationary experiences of the several economies. Dispersion fell in the late 1920s, as a new gold standard was rebuilt, but when this experiment failed dispersion grew again in the 1930s following widespread deflations and various devaluations and depreciations. Dispersion was again high immediately after 1945,

[16] This matters little for interpretation. However, it is reasonable to try the correction. We know that historical price series often show pronounced breaks at some points in time, typically wars and inflations. The accuracy of attempts to bridge price series across such jumps is often questionable, and may be accurate to only an order of magnitude or less. Such effects are common in my sample, especially before and after the two wars, including the bridge over early-1920s hyperinflations in Europe, some 1940s experiences in Europe, and the 1980s in Latin America (Figure 2 shows occasional spikes where severe inflations may have distorted annual averaging too). Allowing the real exchange rate to be rebased (take a redefined mean) before and after such shocks is one way to correct the dispersion measure for such problems.

but fell steadily during the postwar era, both during and after the Bretton Woods regime. Convergence toward equilibrium was tentative during the less-than-strict early years of Bretton Woods before 1960, when pervasive capital controls were the norm. Thereafter, in the (brief) heyday of the postwar fixed-exchange rate system, deviations from equilibrium dipped to historically low levels. Even a modest increase in dispersion in the floating rate era cannot undo the conclusion that from the 1960s to the present we have seen a degree of integration (measured by PPP) not known since the end of the classical gold standard in 1914.

6 Conclusion

This paper has tested a variety of PPP criteria on historical data for twenty countries covering the period after 1880. Overall, the stringent requirements for the Strong PPP criterion cannot be met, at least not across the entire sample of countries. In the long run the real exchange rate is not everywhere stationary. On the other hand, once we allow for a looser specification, it is possible to find support for the Weak PPP hypothesis. With these adjustments, residuals from the Weak PPP equation give a measure of deviations from parity. When dispersion is so calculated, the interwar period emerges as an important watershed, marking an era of increased deviations from parity. Within that window, the Depression emerges as the era when deviations from parity reached historic maxima prior to a sustained decline after 1945 under Bretton Woods. We may also note that the final fracture of the gold standard in the 1930s led to sharp deviations from parity, whereas the collapse of the Bretton Woods fixed-exchange-rate system led to no such sustained departures from parity in the 1970s, even as deviations from parity did increase markedly in their short-run (monthly or quarterly) volatility (Isard 1995, chap. 1).

These findings emerge as a mark against the efficiency and resiliency of the international system at the time of the Depression. Such findings also have broader implications for the debate over globalization and convergence (Williamson 1996). Seminal studies by Abramovitz (1986) and Baumol (1986) noted the long-run patterns of convergence in income per capita in a cross-section of countries. These authors used the historical national income data af Maddison (1991 and his previous studies). Abramovitz noted that not only technological convergence, but also "trade and its rivalries" might have played a part in this historical convergence process, the latter including factor mobility. The convergence findings were corroborated by Williamson (1995) for real wages: divergence was pronounced in the interwar period, and rising through the 1930s. Williamson and his collaborators (see Williamson 1996) have linked these patterns to the evolution and integration of global labor and commodity markets. Of course, in a trade-theoretic framework, there may be substantial complementarities between labor and capital mobility, as in the factor price convergence literature, and as historical experience shows (O'Rourke, Taylor, and Williamson 1996).

Thus, the present paper is certainly consistent with a larger literature on factor mobility, and can help us gauge the importance long-run capital mobility to the historical convergence process—certainly, further reserach is called for on this important subject, especially the interwar experience.[17] The interwar system appears as a terrible aberration, as other studies based on interest rate deviations, saving-investment correlations, and other quantitative criteria are beginning to suggest.[18] Complementing these studies, the present paper strengthens the case that the Great Depression stands as the nadir of international capital market integration in the modern era.

[17] A study which highlights capital mobility and divergence for a single case-study is Taylor (1992) for Argentina. For a cross-sectional study which examines labor and capital mobility, and convergence before 1914, see Taylor and Williamson (1994).

[18] See Eichengreen (1990); Lothian (1995); A. M. Taylor (1996); Obstfeld and Taylor (1996).

Data Appendix

Notation

E = exchange rate; PC = consumption price deflator; PY = GDP deflator.

Sources

Bordo, M. D., and L. Jonung. 1996. Monetary Regimes, Inflation, and Monetary Reform: An Essay in Honor of Axel Leijonhufvud. In *Inflation, Institutions, and Information: Essays in Honor of Axel Leijonhufvud*, edited by D. F. Vaz and K. Vellapillai. London: Macmillan.

Bordo, M. D., and H. Rockoff. 1996. The Gold Standard as a "Good Housekeeping Seal of Approval." *Journal of Economic History* 56 (June): 389428.

Bordo, M. D., and A. J. Schwarz. 1996. The Operation of the Specie Standard: Evidence for Core and Peripheral Countries, 1880–1990. In *Historical Perspectives on the Gold Standard: Portugal and the World*, edited by B. J. Eichengreen and J. B. d. Macedo. London: Routledge.

Maddison, A. 1991. *Dynamic Forces in Capitalist Development: A Long-Run Comparative View.* Oxford: Oxford University Press.

Mitchell, B. R. 1992. *International Historical Statistics: Europe, 1750–1988.* New York: Stockton Press.

World Bank. 1994. *World Data 1994: World Bank Indicators on CD-ROM.* Washington, D.C.: The World Bank.

Argentina

E: 1884–1959, Bordo-Schwarz. 1960–92, World Bank.
PC: 1884–1960, Bordo-Schwarz (PY: 1884–1913). 1960–92, World Bank.

Australia

E: 1870–79, Bordo-Rockoff. 1880–83, Bordo-Schwarz. 1984–92, World Bank.
PC: 1870–80, Bordo-Rockoff. 1880–1959, Bordo-Schwarz. 1960–92, World Bank.

Belgium

E: 1880–1940, 1944–91: Bordo-Jonung.
PC: 1870–1989, Maddison.

Brazil

E: 1889–1991, Bordo-Jonung.
PC: 1880–1991, Bordo-Jonung (PY: 1880–1912).

Canada

E: 1870–79, Bordo-Rockoff. 1880–59, Bordo-Schwarz. 1960–92, World Bank.
PC: 1870–80, Bordo-Rockoff. 1880–1959, Bordo-Schwarz. 1960–92, World Bank.

Chile

E: 1895–1990, Bordo-Jonung.
PC: 1913–91, Bordo-Jonung.

Denmark

E: 1880–1986, Bordo-Jonung. 1987–92, World Bank.
PC: 1850–80, Mitchell. 1880–1960, Bordo-Jonung. 1960–92, World Bank.

Finland

E: 1911–89, Bordo-Jonung.
PC: 1880–1992, Bordo-Jonung.

France

E: 1880–1940, 1948–59, Bordo-Jonung. 1960–92, World Bank.
PC: 1880–1913, 1921–38, 1948–60, Bordo-Jonung. 1960–92, World Bank.

Germany

E: 1880–1914, 1924–39, 1948–89, Bordo-Jonung.
PC: 1870–1989, Maddison

Italy

E: 1880–1959, Bordo-Jonung. 1960–92, World Bank.
PC: 1880–1959, Bordo-Jonung. 1960–92, World Bank.

Japan

E: 1880–1990, Bordo-Jonung.
PC: 1950–89, Bordo-Jonung.
PY: 1885–1950, Bordo-Jonung.

Netherlands

E: 1913–41, 1945–91, Bordo-Jonung.
PC: 1870–1989, Maddison.

Norway

E: 1899–1939, 1946–59, Bordo-Jonung. 1960–92, World Bank.
PC: 1880–1939, 1946–59, Bordo-Jonung. 1960–92, World Bank.

Portugal

E: 1890–1991, Bordo-Jonung.
PC: 1880–1988, Bordo-Jonung (PY: 1880–1929).

Spain

E, PC: 1880–1959, Bordo-Schwarz. 1960–92, World Bank.

Sweden

E: 1899–1959, Bordo-Jonung. 1960–92, World Bank.
PC: 1880–1959, Bordo-Jonung. 1960–92, World Bank.

Switzerland

E: 1880–1991, Bordo-Jonung.
PC: 1892–1989, Maddison.

United Kingdom

E, PC: 1870–79, Bordo-Rockoff. 1880–1959, Bordo-Jonung. 1960–92, World Bank.

United States

PC: 1870–79, Bordo-Rockoff. 1880–1959, Bordo-Jonung. 1960–92, World Bank.

References

Abauf, N., and P. Jorion. 1990. Purchasing Power Parity in the Long Run. *Journal of Finance* 45 (March): 157–74.

Abramovitz, M. 1986. Catching Up, Forging Ahead, and Falling Behind. *Journal of Economic History* 46 (June): 385–406.

Balassa, B. 1964. The Purchasing Power Parity Doctrine. *Journal of Political Economy* 72: 584–96.

Baumol, W. 1986. Productivity Growth, Convergence and Welfare: What the Long-Run Data Show. *American Economic Review* 76 (December): 1072–85.

Cassel, G. 1922. *Money and Foreign Exchange After 1914*. New York: Macmillan.

Diebold, F. X., S. Husted, and M. Rush. 1991. Real Exchange Rates under the Gold Standard. *Journal of Political Economy* 99 (December): 1252–71.

Edison, H. J., J. E. Gagnon, and W. R. Melick. 1994. Understanding the Empirical Literature on Purchasing Power Parity in the Long Run: The Post-Bretton Woods Era. International Finance Discussion Papers no. 465, Board of Governors of the Federal Reserve System.

Edison, H. J. 1987. Purchasing Power Parity in the Long Run: A Test of the Dollar/Pound Exchange Rate, 1890–1978. *Journal of Money, Credit and Banking* 19: 376–87.

Eichengreen, B. J. 1990. Trends and Cycles in Foreign Lending. In *Capital Flows in the World Economy*, edited by H. Siebert. Tübingen: Mohr.

Eichengreen, B. J. 1992. *Golden Fetters: The Gold Standard and The Great Depression, 1919–1939.* Oxford: Oxford University Press.

Engel, C. 1996. Long-Run PPP May Not Hold After All. University of Washington (January). Photocopy.

Engle, R., and C. Granger. 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55: 251–276.

Frankel, J. A. 1986. International Capital Mobility and Crowding Out in the U.S. Economy: Imperfect Integration of Financial Markets and Goods Markets? In *How Open is the U.S. Economy?*, edited by R. Hafer. Lexington: Lexington Books.

Frankel, J. A. 1990. Zen and the Art of Modern Macroeconomics: The Search for Perfect Nothingness. In *Monetary Policy for a Volatile Global Economy*, edited by W. Haraf and T. Willet. Washington, D.C.: American Enterprise Institute.

Frankel, J. A., and A. K. Rose. 1996. A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries. *Journal of International*

Economics 40 : 209–224.

Frenkel, J. 1981. The Collapse of Purchasing Power Parity During the 1970s. *European Economic Review* 16: 145–65.

Froot, K. A., and K. Rogoff. 1995. Perspectives on PPP and Long-Run Real Exchange Rates. In *Handbook of International Economics*, vol. 3, edited by G. Grossman and K. Rogoff. Amsterdam: North Holland.

Glen, J. 1992. Real Exchange Rates in the Short, Medium, and Long Run. *Journal of International Economics* 33: 147–66.

Hansen, B. E. 1995. Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power. *Econometric Theory* 11 (December): 1148–1171.

Heston, A. et al. 1994. The Penn World Table, Version 5.6. National Bureau of Economic Research, Cambridge, Mass. (November). Internet <http://nber.harvard.edu>.

Horvath, M. T. K., and M. W. Watson. 1995. Testing for Cointegration When Some of the Cointegrating Vectors Are Prespecified. *Econometric Theory* 11 (December): 984–1014.

Isard, P. 1995. *Exchange Rate Economics.* Cambridge: Cambridge University Press.

Johnson, D. R. 1990. Cointegration, Error Correction, and Purchasing Power Parity Between Canada and the United States. *Canadian Journal of Economics* 23: 839–55.

Kim, H. 1990. Purchasing Power Parity in the Long Run: A Cointegration Approach. *Journal of Money, Credit and Banking* 22: 491–503.

Kremers, J., N. Ericsson, and J. Dolado. 1992. The Power of Cointegration Tests. *Oxford Bulletin of Economics and Statistics* 54: 325–348.

Lee, M. H. 1978. *Purchasing Power Parity.* New York: Marcel Dekker.

Levin, A., and C.-F. Lin. 1992. Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. Discussion Paper no. 92-23, University of California at San Diego, Department of Economics (May).

Lothian, J. R. 1990. A Century Plus of Yen Exchange Rate Behavior. *Japan and the World Economy* 2 : 47–70.

Lothian, J. R. Capital Market Integration and Exchange-Rate Regimes in Historical Perspective. Fordham University, December 1995. Photocopy.

Lothian, J. R., and M. P. Taylor. 1996. Real Exchange Rate Behavior: The Recent Float From the Perspective of the Past Two Centuries. *Journal of Political Economy.*

Maddison, A. 1991. *Dynamic Forces in Capitalist Development: A Long-Run Comparative View.* Oxford: Oxford University Press.

McCloskey, D. N., and J. R. Zecher. 1984. The Success of Purchasing Power Parity: Historical Evidence and its Implications for Macroeconomics. In *A Retrospective on the Classical Gold Standard, 1880–1913*, edited by M. D. Bordo and A. J. Schwartz. Chicago: University of Chicago Press.

Nurkse, R. 1944. *International Currency Experience: Lessons of the Inter-War Period.* Geneva: League of Nations.

Obstfeld, M., and A. M. Taylor. International Capital Market Integration in the Long Run: The Great Depression as a Watershed. University of California at Berkeley, 1996. Photocopy.

O'Connell, P. G. J. 1996. The Overvaluation of Purchasing Power Parity. Harvard University (January). Photocopy.

Officer, L. H. 1982. *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance.* Greenwich, Conn.: JAI Press.

ORourke, K. H., A. M. Taylor, and J. G. Williamson. 1996. Factor Price Convergence in the Late Nineteenth Century. *International Economic Review.* Forthcoming.

Papell, D. H. 1995. Searching for Stationarity: Purchasing Power Parity Under the Current Float. University of Houston (October). Photocopy.

Pedroni, P. Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series with an Application to the PPP Hypothesis. Working Papers in Economics no. 95-013, Indiana University, June 1995.

Samuelson, P. A. 1964. Theoretical Notes on Trade Problems. *Review of Economics and Statistics* 46: 145–64.

Taylor, A. M. 1992. External Dependence, Demographic Burdens and Argentine Economic Decline After the *Belle Époque*. *Journal of Economic History* 52 (December): 907–36.

Taylor, A. M. 1996. International Capital Mobility in History: The Saving-Investment Relationship in the Long Run. Working Paper Series, National Bureau of Economic Research. Forthcoming.

Taylor, A. M., and J. G. Williamson. 1994. Convergence in the Age of Mass Migration. Working Paper Series no. 4711, National Bureau of Economic Research (April).

Taylor, M. P. The Economics of Exchange Rates. *Journal of Economic Literature* 33 (March 1995): 13–47.

Wei, S.-J., and D. C. Parsley. 1995. Purchasing Power Dis-parity During the

Floating Rate Period: Exchange Rate Volatility, Trade Barriers, and Other Culprits. Working Paper Series no. 5032, National Bureau of Economic Research.

Williamson, J. G. 1995. The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses. *Explorations in Economic History* 32 (April): 141–96.

Williamson, J. G. 1996. Globalization, Convergence, and History. *Journal of Economic History* 56 (June): 277–306.

Yeager, L. B. 1976. *International Monetary Relations: Theory, History, and Policy*. 2d ed. New York: Harper and Row.

Table 1
Unit root and cointegration tests for the real exchange rate

	Using q relative to ROW				Using q relative to U.S.				
		DF	DF	ECM		DF	DF	ECM	HW
Country	T	t	ρ	t	T	t	ρ	t	W
Argentina	70	-2.59	-12.15	-2.61*	109	-2.71	-14.22*	-2.71*	13.33*
Australia	70	-2.25	-9.19	-2.44*	114	-2.69	-12.34	-2.54*	7.58
Belgium	70	-1.79	-3.63	-2.39*	102	-3.36*	-19.94*	-3.94*	23.81*
Brazil	70	-0.30	-0.57	-0.73	103	-1.40	4.32	-1.48	3.64
Canada	70	-1.09	-3.02	-3.93*	116	-3.55*	-17.96*	-3.53*	9.55
Chile					78	-1.06	3.47	-1.05	8.84
Denmark	70	-2.78	-12.43	-1.09	113	-1.82	8.53	-1.30	26.81*
Finland					80	-3.68*	-24.87*	-6.21*	62.95*
France	70	-1.72	-6.05	-1.41	94	-2.80	-13.15	-2.80*	23.00*
Germany	70	1.44	5.16	1.40	91	-1.35	5.13	-0.68	17.09*
Italy	70	-2.90	-9.71	-2.77*	113	-3.05*	-18.34*	-1.87	26.30*
Japan	70	0.60	0.79	0.17	99	0.16	0.33	-0.64	13.56*
Netherlands					73	-1.32	4.60	0.13	18.27*
Norway					87	-1.86	7.65	-2.32*	4.43
Portugal	70	-0.62	-0.74	-0.56	99	-1.87	5.30	-1.89	0.53
Spain	70	-3.58*	-16.60*	-4.22*	101	-1.80	7.82	-2.61*	32.72*
Sweden	70	-1.10	-2.66	-1.29	113	-2.17	-10.77*	-2.25*	44.59*
Switzerland	70	3.19	2.54	2.83*	98	1.06	1.89	1.70	18.51*
United Kingdor	70	-1.79	-6.29	-1.64	114	-2.53	-13.90*	-2.48*	15.95*
United States	70	-1.71	-5.32	-5.99*					

Notes:
* significant at the 5% level. T is sample size. q is log real exchange rate. DF: Dickey-Fuller test for unit root; no trend, no lags; see text. ECM: Error-Correction Model test for cointegration; see text. HW: Horvath-Watson test for cointegration; constant term, no trend; see text.

Table 2
Panel tests for mean reversion of the real exchange rate

	Using q relative to ROW						Using q relative to U.S.					
Sample	NT	N	T	R sq.	β	t	NT	N	T	R sq.	β	t
G	336	16	21	.14	-0.25*	(6.74)	315	15	21	.11	-0.18	(5.51)
I	176	16	11	.15	-0.16	(3.22)	165	15	11	.14	-0.23	(3.90)
B	336	16	21	.13	-0.11	(4.80)	315	15	21	.13	-0.11	(4.95)
F	272	16	17	.15	-0.24	(5.96)	255	15	17	.15	-0.25	(6.08)
GI	512	16	32	.04	-0.06	(3.92)	480	15	32	.04	-0.07	(4.14)
IB	512	16	32	.09	-0.10	(5.79)	480	15	32	.09	-0.13	(6.01)
BF	608	16	38	.07	-0.09	(5.04)	570	15	38	.06	-0.08	(4.79)
GIB	848	16	53	.03	-0.04	(4.33)	795	15	53	.04	-0.05	(4.76)
IBF	784	16	49	.05	-0.07	(4.99)	735	15	49	.05	-0.07	(4.97)
GIBF	1120	16	70	.02	-0.03	(3.39)	1050	15	70	.03	-0.04	(4.11)

Notes:
* significant at the 5% level. G=Gold Standard, I=Interwar, B=Bretton Woods, F=Float. Sample is a panel of 16 countries. U.S. is excluded for q relative to U.S. NT is the number of observations, N is individuals, T is periods. Critical values from Levin and Lin, approximately: 5%, t= -6.5.

Table 3
Dispersion of the real exchange rate

Period	N	V(q,N)	V(q,16)	V(qW,16)
1880-1884	11	0.016		
1885-1889	14	0.056		
1890-1894	16	0.094	0.094	
1895-1899	17	0.098	0.105	0.125
1900-1904	17	0.088	0.093	0.111
1905-1909	17	0.089	0.094	0.110
1910-1914	20	0.065	0.077	0.096
1915-1919	18	0.070		
1920-1924	20	0.041	0.051	0.058
1925-1929	20	0.030	0.036	0.044
1930-1934	20	0.018	0.019	0.025
1935-1939	20	0.030	0.032	0.054
1940-1944	13	0.044		
1945-1949	19	0.064		
1950-1954	20	0.088	0.078	0.082
1955-1959	20	0.066	0.047	0.049
1960-1964	20	0.071	0.025	0.026
1965-1969	20	0.044	0.024	0.026
1970-1974	20	0.033	0.038	0.039
1975-1979	20	0.130	0.096	0.101
1980-1984	20	0.152	0.134	0.144
1985-1989	20	0.316	0.239	

Notes:
q (qW) is the log real exchange rate relative to U.S. (ROW). Quinquennial averages of q or qW in all periods. N is the cross-section width. V(q,N) is the full-sample vaiance. V(q,16) and V(qW,16) are for the panel of 16.

Table 4
Error correction model

	Using q relative to U.S.					Using q relative to ROW				
	T	R sq.	q(t-1)	DrUS(t)	rUS(t-1)	T	R sq.	qW(t-1)	DrW(t)	rW(t-1)
Argentina	108	.07	-0.15*	0.21	-0.02	70	.13	-0.22*	0.07	-0.06
			(-2.70)	(0.46)	(-0.60)			(-3.07)	(0.17)	(-1.63)
Australia	113	.11	-0.13*	-0.29	-0.01	70	.20	-0.36*	-0.16	-0.05
			(-2.58)	(-1.54)	(-0.90)			(-3.97)	(-1.11)	(-3.03)
Belgium	101	.16	-0.40*	1.73	0.02	70	.10	-0.10*	0.21	0.01
			(-4.03)	(2.28)	(0.85)			(-2.34)	(1.65)	(0.77)
Brazil	102	.07	-0.11*	-0.07	-0.05	70	.10	-0.12*	-0.22	-0.07
			(-2.50)	(-0.25)	(-2.01)			(-2.22)	(-1.00)	(-2.14)
Canada	115	.16	-0.21*	-0.06	-0.01	70	.53	-0.10*	-0.61	0.00
			(-4.12)	(-0.84)	(-2.07)			(-3.02)	(-8.40)	(0.41)
Chile	77	.07	-0.11*	0.31	-0.11					
			(-2.13)	(0.57)	(-2.16)					
Denmark	112	.12	-0.11*	-0.52	0.04	70	.43	-0.13*	-0.95	0.05
			(-2.39)	(-2.68)	(2.75)			(-2.14)	(-5.78)	(3.48)
Finland	79	.35	-0.58*	2.22	0.01					
			(-6.16)	(4.75)	(0.20)					
France	93	.10	-0.19*	0.09	-0.02	70	.21	-0.34*	0.15	-0.05
			(-2.84)	(0.28)	(-1.11)			(-3.85)	(1.38)	(-3.61)
Germany	90	.04	-0.03	-0.33	0.01	70	.07	0.10*	0.05	-0.01
			(-0.69)	(-1.18)	(0.53)			(1.79)	(0.60)	(-1.78)
Italy	112	.12	-0.10*	-0.79	0.02	70	.15	-0.14*	-0.12	0.00
			(-1.69)	(-2.32)	(1.15)			(-2.77)	(-1.30)	(-0.51)
Japan	98	.05	-0.09*	0.46	0.05	70	.14	-0.13*	0.11	0.08
			(-1.70)	(2.02)	(1.58)			(-2.66)	(0.85)	(2.94)
Netherlands	74	.07	-0.02	-0.45	0.01					
			(-0.27)	(-1.72)	(0.36)					
Norway	86	.06	-0.14*	0.39	0.00					
			(-2.31)	(1.34)	(0.01)					
Portugal	98	.05	-0.08*	0.25	-0.03	70	.01	-0.02	0.05	-0.01
			(-2.17)	(0.79)	(-1.10)			(-0.85)	(0.41)	(-0.65)
Spain	100	.08	-0.13*	0.53	0.01	70	.29	-0.31*	0.35	0.02
			(-2.65)	(1.79)	(0.59)			(-4.36)	(2.53)	(1.35)
Sweden	112	.07	-0.18*	0.23	0.02	70	.04	-0.04	0.08	0.00
			(-2.70)	(1.04)	(1.50)			(-1.02)	(1.09)	(0.09)
Switzerland	97	.05	0.00	-0.29	0.03	70	.14	0.01	0.04	0.02
			(0.02)	(-1.31)	(0.82)			(-0.14)	(0.37)	(0.69)
United Kingdom	113	.06	-0.12*	0.01	0.00	70	.05	-0.11	-0.03	0.00
			(-2.36)	(0.04)	(0.06)			(-1.61)	(-0.29)	(0.39)
United States						70	.84	-0.10*	-0.79	0.02
								(-5.16)	(-17.78)	(4.70)

Notes:
ECM t-test: * significant at 5% level, one-tailed, asymptotic N(0,1). Constant terms not shown.

Table 5
Panel tests for Weak PPP

	Using q relative to ROW				
Sample	N	T	VSTAT	RHOSTAT	TSTAT
G	16	21	47.75	-85.27*	-14.02*
I	16	11	12.92	-38.99*	-9.11
B	16	21	37.83	-94.62*	-14.33*
F	16	16	33.61	-65.63*	-11.94*
GI	16	32	18.50	-155.35*	-17.34*
IB	16	32	24.82	-158.62*	-18.07*
BF	16	38	57.32*	-185.38*	-19.57*
GIB	16	53	27.55	-277.92*	-23.56*
IBF	16	49	51.71*	-248.16*	-22.43*
GIBF	16	70	40.06	-374.88*	-27.43*

Notes:
* significant at the 5% level. G=Gold Standard, I=Interwar, B=Bretton Woods, F=Float. Sample is a panel of 16 countries. N is the number of countries, T is number of periods. All statistics are for the case of demeaned variables.Critical values and definitions from Pedroni (1995). RHOSTAT is the heterogeneous panel rho statistic. TSTAT is the heterogeneous panel rho-t statistic. VSTAT is the heterogeneous panel variance ratio statistic.

Table 6
Panel test for Weak PPP: Cointegrating regression

Sample	GIBF
NT	1168
N	16
T	73

Country	Slope coefficient
Argentina	0.82
Australia	0.86
Belgium	1.20
Brazil	0.49
Canada	0.88
Denmark	1.09
France	0.86
Germany	1.03
Italy	0.96
Japan	1.50
Portugal	0.65
Spain	1.08
Sweden	1.08
Switzerland	1.64
United Kingdom	0.93
United States	0.94

Notes:
See text. Intercept terms omitted.

Table 7
Dispersion of Weak PPP residuals

Period	N	V(qW resids,N)	V(resids,16,R)
1880-1884			
1885-1889			
1890-1894	15	0.0191	
1895-1899	16	0.0106	0.0012
1900-1904	16	0.0133	0.0005
1905-1909	16	0.0251	0.0009
1910-1914	16	0.0234	0.0005
1915-1919			
1920-1924	16	0.0631	0.0161
1925-1929	16	0.0359	0.0015
1930-1934	16	0.0494	0.0004
1935-1939	16	0.0758	0.0362
1940-1944			
1945-1949			
1950-1954	16	0.0795	0.0032
1955-1959	16	0.0475	0.0024
1960-1964	16	0.0180	0.0096
1965-1969	16	0.0098	0.0007
1970-1974	16	0.0156	0.0013
1975-1979	16	0.0114	0.0068
1980-1984	16	0.0151	0.0051
1985-1989	15	0.0200	

Notes:
qW is the log real exchange rate relative to ROW. Quinquennial averages of q or qW in all periods. N is the cross-section width. V(qWresids,N) is the full-sample variance for resdiuals from the heterogeneous panel OLS regression. In the last columns, R denotes a rebasing relative to means in each period G-I-B-F. V(qWresids,16,R) is the panel-of-16 variance for resdiuals from the heterogeneous panel OLS regression with dummies added for G-I-B-F periods.

Figure 1: Dispersion of real exchange rate

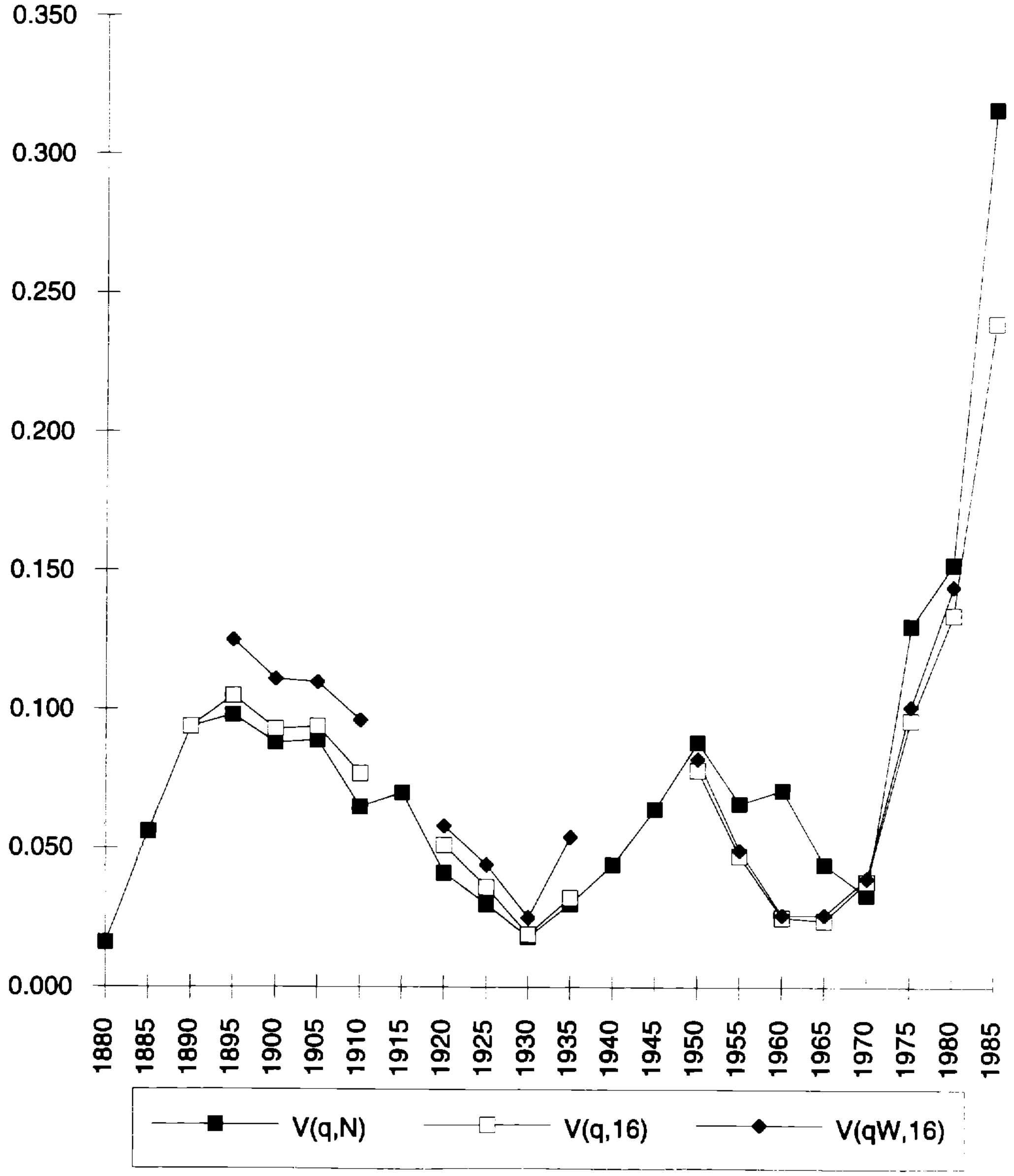

Figure 2: Real Exchange Rates, 1850–1990

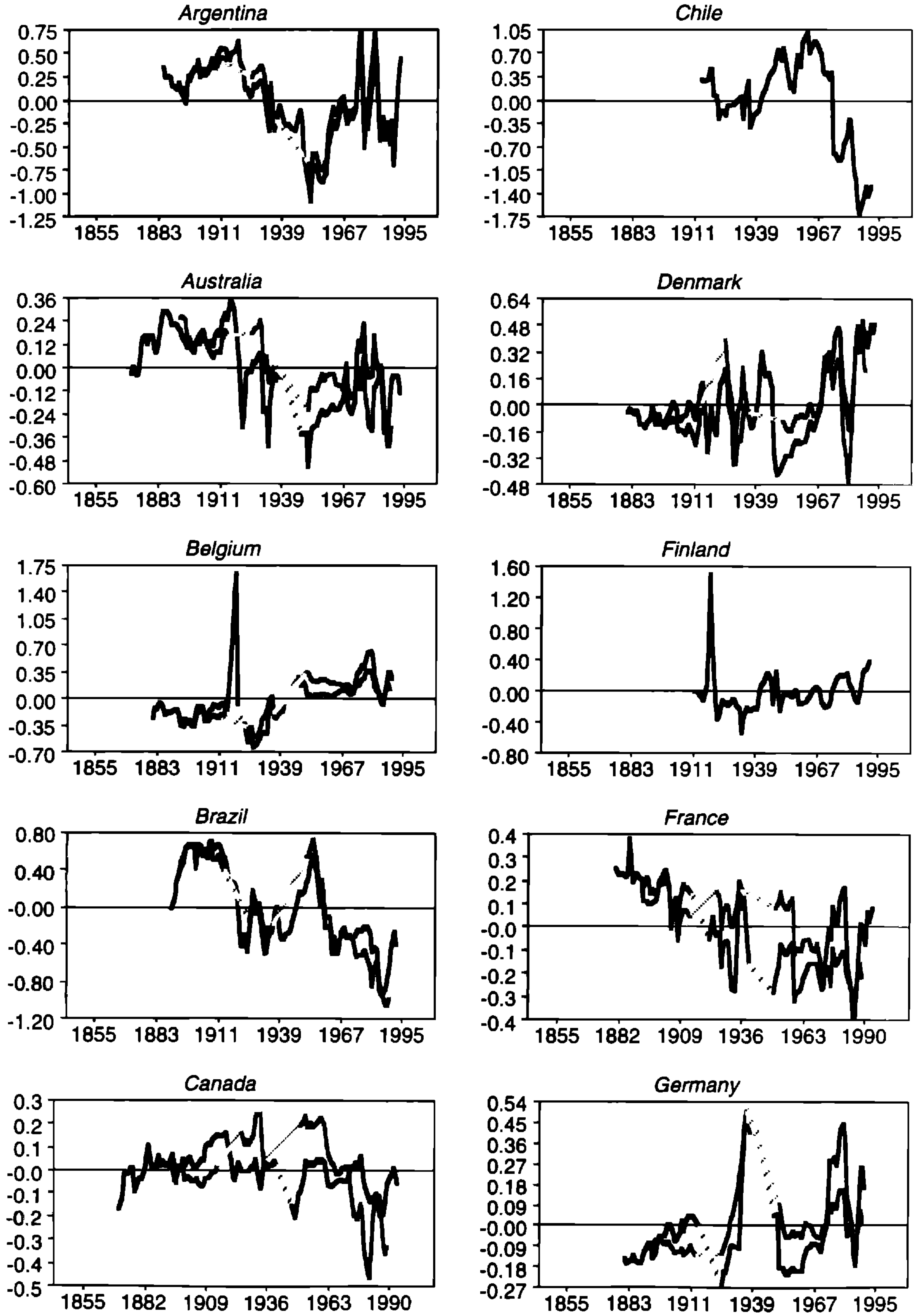

Figure 2 (continued): Real Exchange Rates, 1850–1990

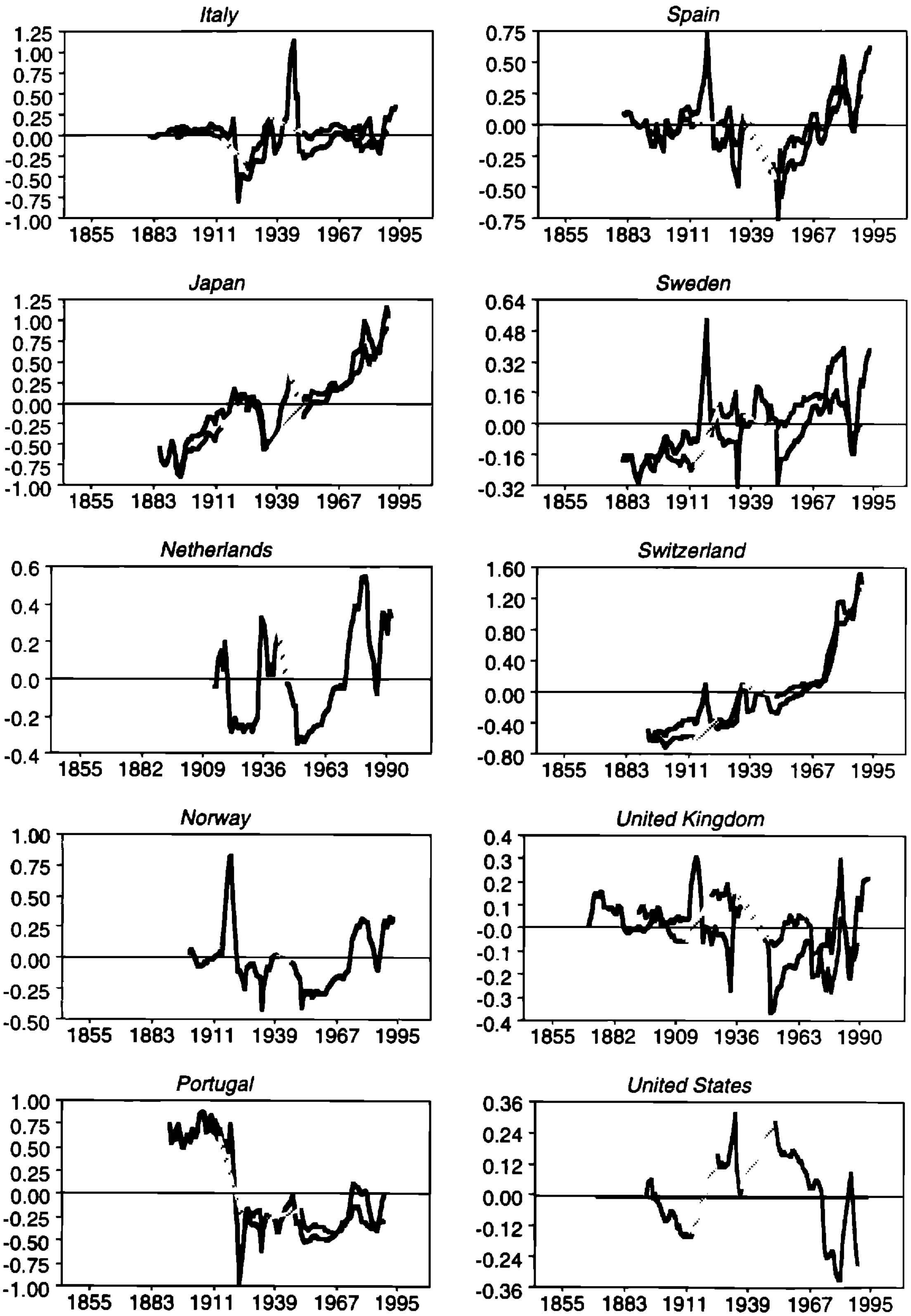

Figure 3: Dispersion of Weak PPP residuals