Real exchange rate behavior

JAMES R LOTHIAN

Graduate School of Business, Fordham University, New York, NY 10023, USA

AND

MARK P TAYLOR

Department of Economics, University of Liverpool, Liverpool L69 3BX, UK and Centre for Economic Policy Research, UK

Using random simulations with artificial data with identical sample characteristics to the long-sample exchange rate data employed by Lothian and Taylor (Lothian, J.R. and Taylor, M.P. (1996). The recent float from the perspective of the past two centuries. Journal of Political Economy 104, 488-509.), we show that standard unit-root tests have extremely low power over sample sizes corresponding to the recent float. The probability of rejecting the null hypothesis when it is false is extremely low with 20 years or even 50 years of data and only reaches an acceptable level over much longer spans. (JEL F31, C15, C22). © 1997 Elsevier Science Ltd. All rights reserved.

Among the puzzles which have emerged concerning the exchange rate over the last 20 years are the following: real exchange rates are hard to distinguish from random walks (or at least martingales); asset-based models of exchange rates, including monetary models, have poor predictive power; nominal exchange rates show much greater variability than fundamental macroeconomic variables; real exchange rates are more variable under a floating rate regime than under a fixed rate regime.1

These are, indeed, among the important stylized facts of the recent float. They are not, however, characteristic of all exchange-rate data. Using a variety of statistical techniques and a much broader data sample than the floating-rate period alone, researchers have now begun to reject the real exchange rate random-walk hypothesis (e.g. Diebold et al., 1991; Lothian and Taylor, 1996). They find instead that real exchange rates revert to equilibrium values over the long-run, and correspondingly, that nominal exchange rates and relative price levels converge, thus reviving the view of purchasing power parity (PPP) as a long-run equilibrium condition.

945
The behavior of real exchange rates under the float may in fact be similar to behavior in other periods and the relative incidence of different types of shocks also largely the same. Researchers may, however, find it difficult to discriminate, since, with less than 20 annual data points, the recent float may simply contain too few low-frequency components for researchers reliably to detect mean-reverting real exchange rate behavior using conventional statistical tests. This point has been made by, among others, Frankel (1986), Huizinga (1987), Abuaf and Jorion (1990), Lothian (1990), Johnson (1993) and Lothian and Taylor (1996). Building on parallel work by the present authors (Lothian and Taylor, 1996), which utilizes 200 years of data on the sterling-dollar exchange rate and nearly that span of data on sterling-franc exchange rates, this paper provides some further evidence to support this view.

In a series of random simulations, we generate artificial data which has the same sample moments as the long-sample data analyzed by Lothian and Taylor (1996). We then generate the empirical power functions of standard unit-root tests applied to these data for samples of 20 years (broadly the length of the recent float) and of 50 years (broadly the length of the post-WWII period) as well as for samples of 100 and 200 years.

Standard tests for mean reversion are shown to have extremely poor power characteristics in the two smaller samples and, for data with sample moments similar to those of the sterling-dollar real exchange rate series, the rejection frequency does not improve significantly with a sample corresponding to 100 years. The implication is that the difficulty in detecting mean reversion in real exchange rates for the recent floating rate period may be due to lack of statistical power in the standard battery of tests.

I. Purchasing power parity: mean reversion in economic thought

Over the past 3 decades, professional thinking on the subject of real exchange rates has shifted radically, and has recently begun to swing back to where it started — itself demonstrating a form of mean reversion. Studies conducted prior to, and in the years immediately following, the move to floating rates generally supported the idea of a tolerably stable long-run real exchange rate. Monetary History of the United States by Friedman and Schwartz (1963) is a particularly prominent example. Although these authors uncover sizable variations in the real exchange rate over various sub-periods, Friedman and Schwartz nevertheless remained impressed by its relative stability over the sample period as a whole, as the following passage clearly indicates (pp. 678–679):

One striking example of the stability of basic economic relations is the stability of relative prices in the United States and Great Britain adjusted for changes in the exchange rate between the dollar and the pound. We have a reasonably continuous series from 1871 on. In the 79 years from 1871 to 1949, vast changes occurred in the economic structure and development of the United States, the place of Britain in the world economy, the internal monetary structures of both the United States and Great Britain, and the international monetary arrangements linking them. Yet despite these changes, despite two world wars and despite the statistical errors in the price-index numbers, the adjusted ratio on the base that makes 1929 = 100 was between 84 and 111 in all but one of the 79 years.
This view, moreover, was not at all atypical. Gaillot (1970, p. 353), in a study using quinquennially averaged data for eight countries over the period 1900–1967, summarized his results as lending support to purchasing power parity as a ‘long-term hypothesis of international economics’ in that they ‘show(ed) the often dramatic effect of non-monetary factors in individual time periods, but at the same time demonstrate(d) the temporary nature of most of these aberrations’.

Some proponents of the monetary approach to exchange rates (originally the monetary approach to the balance of payments) took the argument several steps further. Using the law of one price combined with market efficiency as theoretical underpinnings, they posited a PPP relationship that held over the short-run as well as the long-run (see, e.g. McCloskey and Zecher, 1976).

The examination of the interwar float by Frenkel (1978) likewise reported results broadly favorable to PPP. While eschewing the notion that PPP should be viewed as a theory of exchange rate determination. Frenkel (1978, p. 188) suggested that ‘Its main usefulness is in providing a guide to the general trend of exchange rates’.

By the mid-1980s, in contrast, the majority of researchers had reached very nearly the opposite conclusion. The role of PPP, not only as a rule-of-thumb predictive model but also as an equilibrium condition, became increasingly questioned — PPP was seen to have ‘collapsed’ (Frenkel, 1981a). Krugman (1978), for example, writes: ‘There is some evidence then that there is more to exchange rates than PPP. This evidence is that the deviations of exchange rates from PPP are large, fairly persistent, and seem to be larger in countries with an unstable monetary policy’.2

One of the major reasons for this shift in sentiment was the widespread finding that real exchange rates under the float could be characterized statistically as random walks (e.g. Roll, 1979; Frenkel, 1981a,b; Adler and Lehman, 1983; Darby, 1983). It was confirmed by a subsequent series of studies that applied unit-root tests to real exchange rates and related tests of cointegration between their nominal-exchange-rates and price-level components.3 Reinforcing the change in views about real-exchange-rate stability and PPP were the well-known results of Meese and Rogoff (1983, 1984) that nominal exchange rates could be predicted better by a naive random-walk model than by reduced-form asset market models, as well as the observed shift in the variability of both nominal and real exchange rates under the float. Real exchange rates — particularly those for the US dollar and the pound sterling — showed substantially greater variability under the float than under the previous fixed exchange rate regime. Correspondingly, nominal exchange rates showed much greater variability than important macroeconomic fundamentals such as price levels and real incomes.4

Taken as a group, these findings led to the rather widespread belief that PPP was of little use empirically and that the real exchange rate was highly unstable.

The results of more recent studies, however, are at variance with these conclusions. These studies have been of several types. One group has examined long-term, and, hence, largely pre-float data for the major industrial countries
using tests of cointegration and unit–root tests. (e.g. Lothian and Taylor, 1996). Another has examined shorter-term data for episodes and countries other than the major industrial countries under the float (e.g. Taylor and McMahon, 1988). Both sets of studies find some evidence of reversion of real exchange rates to equilibrium values of one sort or another or, correspondingly, cointegration between price levels and nominal exchange rates.

A number of authors have suggested that the difference between the results of cointegration and unit–root tests with long-term time series and typical results for the float alone may simply be a reflection of the paucity of observations for the latter relative to the degree of variation in the data. In the floating-rate period it may therefore be extremely difficult to distinguish statistically between unit-root and near-unit-root behavior.

One bit of evidence consistent with this conjecture is provided by a third set of studies that use more powerful statistical techniques and expanded data sets to reassess behavior under floating rates. These include the application of conventional tests to multi-country panel data (e.g. Jorion and Sweeney, 1996); the estimation of separate time-varying permanent and transitory components of major-currency real exchange rates (Evans and Lothian, 1993); and tests of long-horizon predictability of major-currency nominal exchange rates, as in Mark (1995). The findings of these studies, in the main, are more favorable to PPP than those of earlier investigations, a result to be expected if paucity of data and low test power had indeed been playing the roles ascribed to them. The results we report below provide a further perspective on this issue.

II. Mean reversion of the real exchange rate

We define the real exchange rate in terms of the logarithmic deviation from purchasing power parity (PPP):

\[ q_t = \ln s_t - (\ln p^*_t - \ln p_t) \]

where \( q_t \) is the logarithm of the real exchange rate, \( s_t \) is the logarithm of the nominal rate (the price in foreign currency of domestic currency), \( p^*_t \) and \( p_t \) are logarithms of the foreign and domestic price indices, respectively, and \( t \) is a time subscript. If (relative) PPP holds continuously, \( q_t \) will be a (non-zero) constant reflecting difference in the units of measurement of \( s_t \) and \( p^*_t - p_t \). There is, however, little evidence of this being the case.

The important question empirically, therefore, is the extent to which PPP holds in the long-run. In dealing with this question, we find it useful conceptually to view \( q_t \) as made up of two components, a long-run equilibrium real exchange rate, \( \bar{q}_t \), and the deviation of \( q_t \) from that equilibrium level:

\[ \bar{q}_t = q_t + (\bar{q}_t - \bar{q}_t) \]

where \( \bar{q}_t = \bar{s}_t - (\bar{p}^*_t - \bar{p}_t) \), and a bar over a variable denotes a long-run value.

If (relative) PPP holds in this long-run context, \( \bar{q}_t \) will be constant and \( q_t \) will ultimately converge to this value, which in turn implies convergence of \( s_t \), \( p_t \), and \( p^*_t \) to their equilibrium values. Over the short-run, however, \( q_t \) need not, and empirically generally does not, equal \( \bar{q}_t \). Divergences will exist so long as
$s_i$, $p_i$, and $p_i^*$ diverge from their long-run equilibria. As a result, tests of long-run PPP have increasingly focused on the error process followed by $q_t$, and in particular whether $q_t$ contains a unit root, or does in fact show convergence to some stable value.

Lothian and Taylor, 1996 demonstrate that sterling real exchange rates against the franc and the dollar over the past 200 years appear to be adequately characterized as realizations from stationary AR(1) processes. Their sample consists of annual observations of dollar–sterling and dollar–franc exchange rates and wholesale price indices of France, the United States and the United Kingdom. In the case of the latter two countries these data span the full 2 centuries 1791–1990, and in the case of France, the 188 years 1803–1990.

Interestingly, the Lothian–Taylor estimation results again indicate a higher degree of persistence in the sterling–dollar real exchange rate than in the franc–dollar real rate. The estimated first-order autocorrelation coefficient for dollar–sterling is 0.887 while for franc–sterling it is 0.776 (Lothian and Taylor, 1996). These estimates indicate that shocks to the real exchange rate are corrected at the rate of some 23% per annum for franc–sterling but at only some 11% per annum for dollar–sterling, implying a half-life of real exchange rate shocks of about 6 years for dollar–sterling and 3 years for franc–sterling. This may be reflecting the larger role which France has traditionally played as a trading partner for the United Kingdom and, relatedly, the closer physical proximity of these two countries: both of these factors will have facilitated the scope for commodity arbitrage.

Below we examine by Monte Carlo methods the empirical power of unit-root tests applied to short-sample realizations of stationary processes similar to those estimated by Lothian and Taylor (1996).

### III. The empirical power of tests for a unit root in the real exchange rate

In this section we describe the Monte Carlo experiments that we use to investigate the empirical power of the standard tests for mean reversion which have been applied to real exchange rates.

We constructed these experiments as follows. At each replication, 300 observations were generated from the AR(1) model:

$$q_t = \alpha_0 + \alpha_1 q_{t-1} + u_t,$$

$$u_t \sim IN(0, \sigma^2),$$

$$q_0 = 0,$$

where $\alpha_0$, $\alpha_1$ and $\sigma^2$ are, respectively, the estimated intercept, slope coefficient and squared standard error of the regression taken from Lothian and Taylor (1996) for the full-sample results for dollar–sterling (case A) and franc–sterling (case B). Case A shows a lower speed of mean reversion of the real exchange rate, with a value of the first-order autocorrelation coefficient, $\alpha_1$, of 0.887. Case B displays a higher speed, with $\alpha_1 = 0.776$.8

949
For both cases, we estimate the Dickey–Fuller ($\tau_\mu$) (Fuller, 1976) and modified Dickey–Fuller [$Z(\tau_\mu)$] (Phillips, 1987; Perron, 1988) unit–root test statistics. These statistics were constructed for four sample sizes: 20 (observations 101 through 120), 50 (observations 101 through 150), 100 (observations 101 through 200), and 200 (observations 101 through 300). This procedure was then replicated 5000 times and the proportion of times that the unit root hypothesis was rejected at the 5% significance level was taken as the empirical power of the test for each case and sample size. For the $Z(\tau_\mu)$ statistic, the asymptotic critical value of $-2.86$ (Fuller, 1976) was used, since Phillips (1987) demonstrated that the limiting distribution of this statistic is the same as that of the Dickey–Fuller statistic. For the $\tau_\mu$ statistic, critical values of $-2.88$, $-2.89$, $-2.93$ and $-3.00$ were used for the sample sizes of 300, 100, 50, and 20, respectively (Fuller, 1976).

IV. The empirical power functions

The results of the Monte Carlo experiments are given in Table 1. They indicate that the power of the Dickey–Fuller and modified Dickey–Fuller tests are only around 9% and 11%, respectively, for an autoregressive time series of length 20 years and root of 0.887. Reducing the root to 0.776 increases the power of these tests only slightly, to 13% and 15%, respectively.

For a sample size of 50, the rejection frequency rises only a little for the higher autoregressive root case A (15% and 18% for $\tau_\mu$ and $Z(\tau_\mu)$, respectively), but increases markedly for the lower autoregressive root case B (42% and 48%, respectively). Raising the sample size to 100 years raises the empirical power of both tests to over 95% for case B, but the higher persistence of case A forces the rejection frequencies to remain below 50%. Thus, even with a century of data on the sterling–dollar real exchange rate, we would have less than an even

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Case A: 1st-order autocorrelation = 0.887</th>
<th>Case B: 1st-order autocorrelation = 0.776</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_\mu$</td>
<td>$Z(\tau_\mu)$</td>
</tr>
<tr>
<td>20</td>
<td>8.60</td>
<td>10.98</td>
</tr>
<tr>
<td>50</td>
<td>15.38</td>
<td>18.28</td>
</tr>
<tr>
<td>100</td>
<td>44.67</td>
<td>48.42</td>
</tr>
<tr>
<td>200</td>
<td>96.20</td>
<td>95.40</td>
</tr>
</tbody>
</table>

Notes: $Z(\tau_\mu)$ was constructed using a Newey and West (1987) lag window, and allowing for up to fifth-order serial correlation. The entries indicate the percentage of times the unit root hypothesis was rejected in 5000 replications. Identical random numbers were used across experiments. The simulations were carried out using the RATS econometric package on an IBM RS6000 machine.
chance of rejecting the unit root hypothesis. With a sample size of 200, however, rejection of the false unit root hypothesis is a virtual certainty for case B, and occurs in over 95% of replications for case A.

V. Conclusion

A number of researchers have conjectured that slow adjustment, coupled with the low power of conventional unit-root tests in any but the longest time series may account for the widespread failure of such tests to reject the unit root null hypothesis in data for the float alone.

The results of the Monte Carlo experiments reported in this paper are entirely consistent with this explanation. For series following first-order autoregressive processes similar to those reported in our earlier study (Lothian and Taylor, 1996), the probability of rejecting the unit-root null when it is not true is extremely low with 20 years, and even 50 years data, and only reaches an acceptable level in both instances over much longer spans.

Notes

3. Earlier studies that use data for the float alone include Enders (1988), Taylor (1988) and Mark (1990). All fail to find cointegration between nominal exchange rates and price levels. In contrast, Huizinga (1987) using spectral analysis and correlation analysis, does find some evidence of long-run real exchange rate mean-reversion for a number of major exchange rates, as also do Abuaf and Jorion (1990), who use pooled data for major industrial countries.
5. The catalogue of studies using long-term time series includes Enders (1989), Abuaf and Jorion (1990), Kim (1990), Lothian (1990) and Diebold et al. (1991), and even earlier Frankel (1986) and Edson (1987). Taken as a group, these studies cover most of the European currencies, the Japanese yen and the Canadian and US dollars. The US dollar typically has been the numeraire, although the yen and sterling have also played that role. Most of these studies use tests of cointegration between components of the real exchange rate or related unit-root tests applied to the real exchange rate itself.
8. The full set of parameter values reported by Lothian and Taylor (1996) were as follows:
    Case A: \( \alpha_0 = 0.179, \alpha_1 = 0.887, \sigma^2 = 0.071^2 \);
    Case B: \( \alpha_0 = -0.309, \alpha_1 = 0.776, \sigma^2 = 0.078^2 \).

9. In any case, the small-sample critical values are larger in absolute value and would lead
to even lower rejection frequencies.

10. Increasing the frequency of observation to quarterly or monthly would not increase the
    low-frequency components of the data and so would be unlikely to improve the
    statistical power (Shiller and Perron, 1985).

References


