Nonlinear Mean-Reversion in Real Exchange Rates: Toward a Solution to the Purchasing Power Parity Puzzles
Author(s): Mark P. Taylor, David A. Peel, Lucio Sarno
Source: *International Economic Review*, Vol. 42, No. 4 (Nov., 2001), pp. 1015-1042
Published by: Blackwell Publishing for the Economics Department of the University of Pennsylvania and Institute of Social and Economic Research -- Osaka University
Stable URL: http://www.jstor.org/stable/826983

INTERNATIONAL ECONOMIC REVIEW
Vol. 42, No. 4, November 2001

NONLINEAR MEAN-REVERSION IN REAL EXCHANGE RATES: TOWARD A SOLUTION TO THE PURCHASING POWER PARITY PUZZLES*

BY MARK P. TAYLOR, DAVID A. PEEL, AND LUCIO SARNO[1]

University of Warwick and CEPR, London; University of Cardiff; and University of Oxford and CEPR, London, United Kingdom

We fit nonlinearly mean-reverting models to real dollar exchange rates over the post-Bretton Woods period, consistent with a theoretical literature on transactions costs in international arbitrage. The half lives of real exchange rate shocks, calculated through Monte Carlo integration, imply faster adjustment speeds than hitherto recorded. Monte Carlo simulations reconcile our results with the large empirical literature on unit roots in real exchange rates by showing that when the real exchange rate is nonlinearly mean reverting, standard univariate unit root tests have low power, while multivariate tests have much higher power to reject a false null hypothesis.

1. INTRODUCTION

Purchasing power parity (PPP) is the proposition that national price levels should be equal when expressed in a common currency. Since the real exchange rate is the nominal exchange rate adjusted for relative national price levels, variations in the real exchange rate represent deviations from PPP. Although the term "purchasing power parity" was coined as recently as 80 years ago (Cassel, 1918), it has a very much longer history in economics.[2] While very few contemporary economists would hold that PPP holds continuously in the real world, "most instinctively believe in some variant of purchasing power parity as an anchor for long-run real exchange rates"

* Manuscript received November 1999; revised July 2000.

[1] The research reported in this article was conducted with the aid of a research grant from the Leverhulme Trust, which Taylor gratefully acknowledges. It consolidates research reported in preliminary form in Taylor and Peel (1998) and Taylor and Sarno (1999a, 1999b). The authors are grateful for constructive comments on earlier drafts to an editor of this journal, Francis Diebold, as well as to four anonymous referees. They are also grateful to participants in seminars at which previous versions of the article were presented, including presentations at Columbia University; the Stern School of Business, New York University; the University of Houston; Rice University; the London School of Economics; the London Business School; the University of Oxford; the University of Salerno; the University of Nottingham; the University of Durham; the University of Southampton; the University of Warwick; Brunel University; the Bank of England; the 1999 ASSA Annual Meetings, New York; and the 1999 Annual Conference of the *Society for Nonlinear Dynamics and Econometrics*, New York. The authors alone are responsible for any errors that may remain.

[2] See Officer (1982) for a fascinating and scholarly account of the history of thought on purchasing power parity.

(Rogoff, 1996), and indeed the implication or assumption of much reasoning in international macroeconomics is that some form of PPP holds at least as a long-run relationship.[3] Moreover, estimates of PPP exchange rates are important for practical purposes such as determining the degree of misalignment of the nominal exchange rate and the appropriate policy response, the setting of exchange rate parities, and the international comparison of national income levels. It is not surprising, therefore, that a large empirical literature on PPP has evolved. This literature, at least as it has developed since the late 1980s, has thrown up at least two important puzzles.

A large number of studies of real exchange rates among industrialized countries in the post-Bretton Woods period have found that time-series deviations from PPP appear to be characterized by nonstationary (or more particularly unit-root) behavior for exchange rates, implying the absence of any long-run tendency toward PPP (for recent surveys, see, e.g., Froot and Rogoff, 1996; Rogoff, 1996; Taylor, 1995). This might be thought of as the first PPP puzzle. One potential solution to the first PPP puzzle is by reference to the low power of the statistical tests employed. In particular, since conventional tests for unit roots may have very low power to reject a false null hypothesis with a sample span corresponding to the length of the recent float, i.e., since about 1973 (Frankel, 1986, 1990; Lothian and Taylor, 1997), a number of researchers have sought to increase the power of their tests by increasing the length of the sample period under consideration (e.g., Cheung and Lai, 1993a; Diebold et al., 1991; Lothian and Taylor, 1996), and have in fact been able to find significant evidence of real exchange rate mean reversion. Other researchers have sought to increase test power by using panel unit root tests applied jointly to a number of real exchange rate series over the recent float and, in many of these studies, the unit-root hypothesis is also rejected for groups of real exchange rates (e.g., Frankel and Rose, 1996; Wu, 1996; Flood and Taylor, 1996; Papell, 1998; Taylor and Sarno, 1998).[4]

In our view, however, whether or not the long-span or panel-data studies do in fact solve the first PPP puzzle remains contentious. As far as the long-span studies are concerned, as noted in particular by Frankel and Rose (1996), the long samples required to generate a reasonable level of statistical power with standard univariate unit root tests may be unavailable for many currencies (perhaps thereby generating a "survivorship bias" in tests on the available data—Froot and Rogoff, 1996) and, in any case, may potentially be inappropriate because of differences in real exchange rate behavior both across different historical periods and across different nominal exchange rate regimes (e.g., Baxter and Stockman, 1989; Hegwood and Papell, 1998). As for panel-data studies, one potential problem with panel unit root tests, highlighted by the Monte Carlo evidence of Taylor and Sarno (1998), is that the null hypothesis in such tests is generally that *all* of the series are generated by unit-root processes, so that the probability of rejection of the null hypothesis may be

[3] This is true both of traditional international macroeconomic analysis (e.g., Dornbusch, 1980) and of the "new" open economy macroeconomics based on intertemporal optimizing models (Obstfeld and Rogoff, 1995, 1996).

[4] A third method for increasing the power of unit root tests, by employing univariate tests based on generalized or weighted least squares estimators, has recently been proposed by Cheung and Lai (1998).

quite high when as few as just one of the series under consideration is a realization of a stationary process.

Moreover, even if we take the results of the long-span or panel-data studies as having solved the first PPP puzzle, a second PPP puzzle then arises as follows. Among the long-span and panel-data studies which do report significant mean-reversion of the real exchange rate, there appears to be a consensus that the size of the half-life of deviations from PPP is about three to five years (Rogoff, 1996). If we take as given that real shocks cannot account for the major part of the short-run volatility of real exchange rates (since it seems incredible that shocks to real factors such as tastes and technology could be so volatile) and that nominal shocks can only have strong effects over a time frame in which nominal wages and prices are sticky, then a second PPP puzzle is the apparently high degree of persistence in the real exchange rate (Rogoff, 1996).[5]

We seek to resolve the first and second purchasing power parity puzzles through an investigation of nonlinearities in real exchange rate adjustment toward long-run equilibrium during the post-Bretton Woods period.[6] We provide evidence of nonlinear mean reversion in a number of major real exchange rates, such that real exchange rates behave more like unit root processes the closer they are to long-run equilibrium and, conversely, become more mean reverting the further they are from equilibrium. Moreover, while small shocks to the real exchange rate around equilibrium will be highly persistent, larger shocks mean-revert much faster than the "glacial rates" previously reported for linear models (Rogoff, 1996). Further, we reconcile our findings with the huge literature on unit roots in real exchange rates through Monte Carlo studies and, in particular, demonstrate that when the true data generating process implies nonlinear mean reversion of the real exchange rate, standard univariate unit root tests will have very low power, while multivariate unit root tests will have much higher power to reject a false null hypothesis of unit root behavior.

The rest of the article is structured as follows. In the next section we provide a brief review of the emerging theoretical literature on nonlinearities in real exchange rate adjustment as an additional motivation for our research, while in Section 3 we discuss the class of nonlinear models which we examine and other aspects of our econometric methods. Section 4 describes our data set. In Section 5 we report the results of univariate and multivariate unit root tests applied to the data, as well as describing our multivariate unit root tests as a prelude to the Monte Carlo studies reported in Section 8. In Section 6 we report our nonlinear estimation results and tests for nonlinear mean reversion of the real exchange rate to see if we can find evidence of

[5] Allowing for underlying shifts in the equilibrium dollar-sterling real exchange rate [Harrod–Balassa–Samuelson (HBS) effects] over the past 200 years through the use of nonlinear time trends, Lothian and Taylor (2000) suggest that the half-life of deviations from PPP for this exchange rate may in fact be as low as 2 1/2 years. The evidence in favor of HBS effects in dollar real exchange rates during the recent float, however, seems slight (Engel, 1999).

[6] Michael et al. (1997) investigate nonlinearities in real exchange rates for the 1920s and using the Lothian and Taylor (1996) two-century span of data, although evidence of real exchange rate stability in these data sets is relatively uncontentious (Taylor and McMahon, 1988; Lothian and Taylor, 1996) compared to the post-Bretton Woods period alone (Froot and Rogoff, 1995; Taylor, 1995; Rogoff, 1996).

mean reversion of real exchange rates over the recent floating rate period and resolve the first PPP puzzle. In Section 7 we examine the half-lives implied by the estimated nonlinear models, using Monte Carlo integration, to examine whether our results can shed some light on the second PPP puzzle of very high real exchange rate persistence. Section 8 contains the results of a number of Monte Carlo experiments designed to aid our understanding of the large empirical literature which has tested for unit roots in real exchange rates—including both univariate and multivariate tests—in the light of our nonlinear estimation results. A final section briefly summarizes and concludes. In the Appendix we describe in more detail the methods used to obtain estimates of real exchange rate persistence based on the estimated nonlinear models and the results obtained.

2. THEORETICAL STUDIES OF NONLINEAR REAL EXCHANGE RATE ADJUSTMENT: A BRIEF REVIEW

The idea that there may be nonlinearities in real exchange rate adjustment dates at least from Heckscher (1916), who suggested that there may be significant deviations from the law of one price due to international transactions costs between spatially separated markets. A similar viewpoint can be discerned in the writings of Cassel (e.g., Cassel, 1922) and, to a greater or lesser extent, in other earlier writers (Officer, 1982). More recently, a number of authors have developed theoretical models of nonlinear real exchange rate adjustment arising from transactions costs in international arbitrage (e.g., Benninga and Protopapadakis, 1988; Williams and Wright, 1991; Dumas, 1992; Sercu et al., 1995; O'Connell, 1997; Ohanian and Stockman, 1997). In most of these models, proportional or "iceberg" transport costs ("iceberg" because a fraction of goods are presumed to "melt" when shipped) create a band for the real exchange rate within which the marginal cost of arbitrage exceeds the marginal benefit. Assuming instantaneous goods arbitrage at the edges of the band then typically implies that the thresholds become reflecting barriers.

Drawing on recent work on the theory of investment under uncertainty, some of these studies show that the thresholds should be interpreted more broadly than as simply reflecting shipping costs and trade barriers *per se*, but also as resulting from the sunk costs of international arbitrage and the resulting tendency for traders to wait for sufficiently large arbitrage opportunities to open up before entering the market (see, in particular, Dumas, 1992; also Dixit, 1989; Krugman, 1989).

O'Connell and Wei (1997) extend the iceberg model to allow for fixed as well as proportional costs of arbitrage. This results in a two-threshold model where the real exchange rate is reset by arbitrage to an upper or lower inner threshold whenever it hits the corresponding outer threshold. Intuitively, arbitrage will be heavy once it is profitable enough to outweigh the initial fixed cost but will stop short of returning the real rate to the PPP level because of the proportional arbitrage costs. Coleman (1995) argues that the assumption of instantaneous trade should be replaced with the presumption that it takes time to ship goods. In this model, transport costs again create a band of no arbitrage for the real exchange rate, but the exchange rate can stray beyond the thresholds. Once beyond the upper or lower threshold, the real rate becomes increasingly mean-reverting with the distance from the threshold. Within the transactions costs band, when no trade takes place, the process is divergent so that the exchange rate spends most of the time away from parity.

Overall, these models suggest that the exchange rate will become increasingly mean-reverting with the size of the deviation from the equilibrium level. In some models the jump to mean-reverting behavior is sudden, while in others it is smooth, and Dumas (1994) suggests that even in the former case, time aggregation will tend to smooth the transition between regimes. Moreover, if the real exchange rate is measured using price indices made up of goods prices each with a different size of international arbitrage costs, one would expect adjustment of the overall real exchange rate to be smooth rather than discontinuous.

3. MODELING NONLINEAR MEAN-REVERSION

The real exchange rate, q_t, may be expressed in logarithmic form as

$$q_t \equiv s_t - p_t + p_t^* \tag{1}$$

where s_t is the logarithm of the nominal exchange rate (domestic price of foreign currency), and p_t and p_t^* denote the logarithms of the domestic and foreign price levels, respectively. The real exchange rate may thus be interpreted as a measure of the deviation from PPP. In the procedures conventionally applied to test for long-run PPP, the null hypothesis is usually that the process generating the real exchange rate series has a unit root, while the alternative hypothesis is that all of the roots of the process lie within the unit circle. Thus, the maintained hypothesis in the conventional framework assumes a linear autoregressive process for the real exchange rate, which means that adjustment is both continuous and of constant speed, regardless of the size of the deviation from PPP. As noted above, however, the presence of transactions costs may imply a nonlinear process which has important implications for the conventional unit root tests of long-run PPP. Some empirical evidence of the effect of transactions costs on tests of PPP is provided by Davutyan and Pippenger (1990). More recently, Obstfeld and Taylor (1997) have investigated the nonlinear nature of the adjustment process in terms of a threshold autoregressive (TAR) model (Tong, 1990). The TAR model allows for a transactions costs band within which no adjustment takes place—so that deviations from PPP may exhibit unit root behavior—while outside of the band the process switches abruptly to become stationary autoregressive. While discrete switching of this kind may be appropriate when considering the effects of arbitrage on disaggregated goods prices (Obstfeld and Taylor, 1997), discrete adjustment of the aggregate real exchange rate would clearly be most appropriate only when firms and traded goods are identical. Moreover, many of the theoretical studies discussed above suggest that smooth rather than discrete adjustment may be more appropriate in the presence of proportional transactions costs and, as suggested by Teräsvirta (1994), Dumas (1994), and Bertola and Caballero (1990), time aggregation and nonsynchronous adjustment by heterogeneous agents is likely to result in smooth aggregate regime switching.[7]

[7] We did, however, attempt to test for TAR adjustment in our empirical work—see footnote 25.

An alternative characterization of nonlinear adjustment, which allows for smooth rather than discrete adjustment, is in terms of a smooth transition autoregressive (STAR) model (Granger and Teräsvirta, 1993). In the STAR model, adjustment takes place in every period but the speed of adjustment varies with the extent of the deviation from parity. A STAR model may be written

$$(2) \qquad [q_t - \mu] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu] + \left[\sum_{j=1}^{p} \beta_j^* [q_{t-j} - \mu] \right] \Phi[\theta; q_{t-d} - \mu] + \varepsilon_t$$

where $\{q_t\}$ is a stationary and ergodic process, $\varepsilon_t \sim \text{iid}(0, \sigma^2)$, and $(\theta\, \mu) \in \{\Re^+ \times \Re\}$, where $\Re$ denotes the real line $(-\infty, \infty)$ and $\Re^+$ the positive real line $(0, \infty)$. The transition function $\Phi[\theta; q_{t-d} - \mu]$ determines the degree of mean-reversion and is itself governed by the parameter θ, which effectively determines the speed of mean-reversion, and the parameter μ which is the equilibrium level of $\{q_t\}$. A simple transition function suggested by Granger and Teräsvirta (1993) is the exponential function

$$(3) \qquad \Phi[\theta; q_{t-d} - \mu] = 1 - \exp\left[-\theta^2 [q_{t-d} - \mu]^2\right]$$

in which case (2) would be termed an exponential STAR or ESTAR model. The exponential transition function is bounded between zero and unity, $\Phi\colon \Re \to [0, 1]$, has the properties $\Phi[0] = 0$ and $\lim_{x \to \pm\infty} \Phi[x] = 1$, and is symmetrically inverse–bell-shaped around zero. These properties of the ESTAR model are attractive in the present modelling context because they allow a smooth transition between regimes and symmetric adjustment of the real exchange rate for deviations above and below the equilibrium level.[8] The transition parameter θ determines the speed of transition between the two extreme regimes, with lower absolute values of θ implying slower transition. The inner regime corresponds to $q_{t-d} = \mu$, when $\Phi = 0$ and (2) becomes a linear AR(p) model:

$$(4) \qquad [q_{t-d} - \mu] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu] + \varepsilon_t$$

[8] The class of nonlinear models is infinite, and we have chosen to concentrate on the ESTAR formulation primarily because of these attractive properties, its relative simplicity, and the large amount of previous research on the estimation of STAR models. An alternative approach might be to consider, for example, a model in which the threshold is triggered by the nature (sign and size) of the shocks impinging upon the system, such as in Elwood's (1998) asymmetric moving average model of real output movements; this did not seem to us particularly attractive in the present context, however, since arbitrage activities are likely to be triggered by the level of the real exchange rate itself rather than by the nature of shocks impinging upon it. Similarly, we considered the floor/ceiling model of Pesaran and Potter (1997) as more applicable in a context where discrete rather than smooth adjustment is expected and where some asymmetry in adjustment is expected. An extended form of Hamilton's (1989) Markov-switching model, in which the transition probability is a function of, *inter alia*, the lagged deviation of the real exchange rate from its equilibrium level (Diebold et al., 1994) is, however, currently under investigation by one of the present authors.

The outer regime corresponds, for a given θ, to $\lim_{[q(t-d)-\mu]\to\pm\infty} \Phi[\theta; q_{t-d} - \mu]$, where (2) becomes a different AR(p) model,

$$[q_{t-d} - \mu] = \sum_{j=1}^{p}(\beta_j + \beta_j^*)[q_{t-j} - \mu] + \varepsilon_t \tag{5}$$

with a correspondingly different speed of mean reversion so long as $\beta_j^* \neq 0$ for at least one value of j.

Granger and Teräsvirta (1993) also suggest the logistic function as a plausible transition function for some applications, resulting in a logistic STAR or LSTAR model. Since, however, the LSTAR model implies asymmetric behavior of q according to whether it is above or below the equilibrium level, we regard that model, *a priori*, as inappropriate for modelling real exchange rate movements. That is to say, it is hard to think of economic reasons why the speed of adjustment of the real exchange rate should vary according to whether the dollar is overvalued or undervalued, especially if one is thinking of goods arbitrage as ultimately driving the impetus toward the long-run equilibrium and one is dealing with major dollar real exchange rates against the currencies of other developed industrialized countries. We do, however, test for nonlinearities arising from the LSTAR formulation as a test of specification of our estimated models.[9]

It is also instructive to reparameterize the STAR model (2) as

$$\begin{aligned}\Delta q_t = {} & \alpha + \rho q_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta q_{t-j} \\ & + \left\{\alpha^* + \rho^* q_{t-1} + \sum_{j=1}^{p-1} \phi_j^* \Delta q_{t-j}\right\} \Phi[\theta; q_{t-d}] + \varepsilon_t\end{aligned} \tag{6}$$

where $\Delta q_{t-j} \equiv q_{t-j} - q_{t-j-1}$. In this form, the crucial parameters are ρ and ρ^*. Our discussion of the effect of transactions costs in the previous section suggests that the larger the deviation from PPP, the stronger the tendency to move back to equilibrium.[10] This implies that while $\rho \geq 0$ is admissible, we must have $\rho^* < 0$ and $(\rho + \rho^*) < 0$. That is, for small deviations q_t may be characterized by unit root or even explosive behavior, but for large deviations the process is mean reverting. This analysis has implications for the conventional test for a unit root in the real exchange

[9] Note that we have also chosen the "transition variable" driving the speed of transition as the d-lagged deviation of q from the equilibrium level μ. A possible alternative transition variable would be the d-lagged change in q. This does not seem to us to be particularly attractive in the case of real exchange rate modelling, however, since it would imply, for example, that if the real exchange rate were to be stuck at a level away from equilibrium for d periods, then it would tend to get stuck there indefinitely (the speed of mean reversion would be very low) until there were a sizeable exogenous shock.

[10] Note that in Equation (6) we have written the intercepts α and α^* in unrestrictive form for expositional simplicity, since this corresponds to the standard form in which the augmented Dickey–Fuller regression is usually estimated—i.e., Equation (7). In our empirical work, however, the STAR models were estimated as in Equation (2).

rate process, which is based on a linear AR(p) model, written below as an augmented Dickey–Fuller regression:

$$\Delta q_t = \alpha' + \rho' q_{t-1} + \sum_{j=1}^{p-1} \phi'_j \Delta q_{t-j} + \varepsilon_t \tag{7}$$

Assuming that the true process for q_t is given by the nonlinear model (6), estimates of the parameter ρ' in (7) will tend to lie between ρ and $(\rho + \rho^*)$, depending upon the distribution of observed deviations from the equilibrium level μ. Hence, the null hypothesis $H_0 : \rho' = 0$ (a single unit root) may not be rejected against the stationary linear alternative hypothesis $H_1 : \rho' < 0$, even though the true nonlinear process is globally stable with $(\rho + \rho^*) < 0$. Thus, failure to reject the unit root hypothesis on the basis of a linear model does not necessarily invalidate long-run PPP. This is an issue to which we will return in our Monte Carlo analysis of unit root tests in Section 8.

In empirical applications, Granger and Teräsvirta (1993) and Teräsvirta (1994) suggest choosing the order of the autoregression, p, through inspection of the partial autocorrelation function, PACF; the PACF is to be preferred to the use of an information criterion since it is well known that the latter may bias the chosen order of the autoregression toward low values, and any remaining serial correlation may affect the power of subsequent linearity tests. Granger and Teräsvirta (1993) and Teräsvirta (1994) then suggest applying a sequence of linearity tests to artificial regressions similar in form to (2) but with the second term on the right-hand side replaced by cross products of q_{t-j} and first, second, and third powers of q_{t-d}, for various values of d.[11] These can be interpreted as second- or third-order Taylor series expansions of (2). This allows detection of general nonlinearity through the significance of the higher-order terms, with the value of d selected as that giving the largest value of the test statistic. The tests can also be used to discriminate between ESTAR and LSTAR formulations, since third-order terms disappear in the Taylor series expansion of the ESTAR transition function.[12] We term this method of selecting the order of d and choosing whether an ESTAR or LSTAR formulation is appropriate the Teräsvirta Rule. In the Monte Carlo study of Teräsvirta (1994), the Teräsvirta Rule worked well in selecting d and also in discriminating between ESTAR and LSTAR unless, understandably, the two models are close substitutes—i.e. when most of the observations lie above the equilibrium level μ so that only one half of the inverse-bell shape of the ESTAR transition function is relevant and is well approximated by a logistic curve.

A feature of the STAR family is that deviations from the equilibrium level only generate increasing mean reversion with a delay—of d periods in fact in Equation (2). Of course, all econometric models are only an approximation to vastly more sophisticated real world data generating processes, so that this feature should not worry us unduly. Nevertheless, in the modeling of real exchange rate behavior, economic intuition suggests a presumption in favor of smaller values of the delay parameter d

[11] See also Luukkonen et al. (1988).

[12] Intuitively, this is because the inverse-bell shape of the ESTAR transition function (3) is better approximated by a quadratic than by a cubic curve.

rather than larger values, in that it is hard to imagine why there should be very long lags before the real exchange rate begins to adjust in response to a shock. Given that our economic intuition suggests, in the present application, a presumption in favor of an ESTAR over an LSTAR formulation and a low value for d, we also employed a complementary empirical method for selecting the appropriate nonlinear model as follows. We did this for two reasons: first, as a robustness check on the Teräsvirta Rule and, second, because the Taylor series expansion results underpinning the Teräsvirta Rule require a stationarity assumption which may be violated if, in fact, the real exchange rate is actually generated by a unit root process.[13] Our complementary model selection procedure is as follows.

The order of the autoregression, p, is chosen through inspection of the PACF as before. An ESTAR model is then estimated with d set equal to unity. Lagrange multiplier (LM) tests of the form suggested by Eitrheim and Teräsvirta (1996) are then applied to the residuals of the estimated equation to test the hypothesis of no remaining nonlinearity of the ESTAR type for a range of values of d greater than unity. If significant remaining nonlinearity is detected, d is increased, the model is re-estimated, and the LM tests are applied again. Once a model is arrived at for which the LM tests for no remaining nonlinearity are all insignificant at the chosen significance level, then the specification search ceases. The residuals of that estimated equation are then tested for no remaining nonlinearity of the LSTAR type, again using LM tests of the type described by Eitrheim and Teräsvirta (1996), as a test of specification.

As a further check, we also estimated the delay parameter d along with the other model parameters, by nonlinear least squares through a grid search procedure.

4. DATA

The data set comprises monthly observations on consumer price indices for the US, the UK, Germany, France, and Japan and end-of-period spot exchange rates for UK sterling, German mark, French franc, and Japanese yen against the US dollar. All data cover the sample period from 1973M01 to 1996M12 and are taken from the International Monetary Fund's *International Financial Statistics* database. Real exchange rate series were constructed with these data in logarithmic form as in Equation (1), with s_t taken as the logarithm of the dollar price of currency, p_t as the logarithm of the US consumer price level, and p_t^* as the logarithm of the consumer price level of the relevant country. The (log) real exchange rate series were then normalized on the beginning of the sample so that $q(1973M01) = 0$.

5. UNIT ROOT TESTS[14]

5.1. *Univariate Unit Root Tests.* As a preliminary exercise, we tested for unit root behavior of each of the (log) real exchange rate series by calculating standard Dickey–Fuller test statistics (Table 1). In each case a simple, nonaugmented Dickey–Fuller statistic appeared adequate in that no residual autocorrelation was evident in

[13] The authors are grateful to an anonymous referee for pointing this out.

[14] In all statistical tests executed in this and subsequent sections we use a five percent nominal level of significance, unless specified otherwise.

TABLE 1
UNIVARIATE LINEAR UNIT ROOT TESTS

	q	$q^{(\tau)}$	Δq	$\Delta^2 q$
UK	−2.2324	−2.3784	−15.3296	−27.3896
Germany	−1.7217	−1.7313	−16.6078	−29.7816
France	−1.7467	−1.7539	−17.1910	−30.1339
Japan	−1.5261	−1.9470	−16.3460	−30.1529

NOTES: Statistics are univariate (nonaugmented) Dickey–Fuller test statistics for the null hypothesis of a unit root process; q denotes the log-level of the real exchange rate; the (τ) superscript indicates that a linear trend was also included in the Dickey–Fuller regression; Δ is the first-difference operator. The critical value at the five percent level of significance is −2.88 to two decimal places if no linear trend is included in the regression, and −3.43 if a linear trend is also included in the regression (Fuller, 1976; MacKinnon, 1991).

the auxiliary regressions.[15] In keeping with the very large number of studies of unit root behavior for these exchange rates, we were in each case unable to reject the unit root null hypothesis at conventional nominal levels of significance. On the other hand, putting the series into first-difference form did appear to induce stationarity. Taken alone, these results would confirm the often-recorded result that dollar real exchange rates appear to be generated by processes with a single unit root—a finding which has the important corollary that long-run PPP is violated during the recent float across the major industrialized countries (e.g., Taylor, 1988; Mark, 1990).

5.2. *Multivariate Unit Root Tests.*

5.2.1. *MADF and MADFτ tests.* In keeping with the more recent empirical literature on mean-reversion in real exchange rates, we also applied two linear multivariate unit root tests to the data. The first test, the multivariate augmented Dickey–Fuller test or MADF, was introduced by Taylor and Sarno (1998) and is a multivariate analogue of the standard, single-equation augmented Dickey–Fuller (ADF) test. As such, its properties are essentially similar to other panel unit root tests recently applied by a number of authors.[16] Essentially, the MADF statistic is constructed by estimating autoregressions for each of the real exchange rates jointly, using a seemingly unrelated regressions estimator in order to take account of contemporaneous correlations among the disturbances (O'Connell, 1998), and then testing the joint null hypothesis that the autoregressive coefficients sum to unity in each equation by a standard Wald test. The resulting Wald statistic is the MADF statistic.[17] As the small-sample distribution of the MADF statistic is unknown because of the presence of a unit root under the null hypothesis, its finite-sample empirical distribution must be calculated by Monte Carlo simulation.

[15] Moreover, higher-order augmented Dickey-Fuller tests yielded results qualitatively identical to those reported in Table 1.
[16] In particular, the MADF can be seen as a generalization of the multivariate unit root test suggested by Abuaf and Jorion (1990)—see Taylor and Sarno (1998).
[17] See Taylor and Sarno (1998: pages 287–288) for the exact formula for the MADF statistic.

TABLE 2

ESTIMATED CRITICAL VALUES AT THE FIVE PERCENT LEVEL FOR THE MULTIVARIATE UNIT ROOT TESTS

(a) *MADF and MADFτ tests*						
	$T = 50$	$T = 75$	$T = 100$	$T = 200$	$T = 288$	$T = 300$
MADF	18.0717	16.1257	14.7902	13.0286	12.6322	12.5251
MADFτ	23.3491	20.0596	19.9665	16.6761	16.3693	16.1589
(b) *JLR test*						
	$T = 50$	$T = 75$	$T = 100$	$T = 200$	$T = 288$	$T = 300$
	4.5816	4.1472	3.9966	3.9324	3.9113	3.9018

NOTES: *Panel* (a): The empirical critical values correspond to a first-order, four-variable system calibrated on the real exchange rate data. Each experiment involved 5000 replications and the critical value was taken as the 95th percentile. T is the sample size in the replications. *Panel* (b): Using a first-order, four-variable system calibrated on the real exchange rate data, for each sample size four experiments were performed in which, respectively, all four, the first three, the first two, and the first of the four autoregressive equations in the data generating process had a unit root while the remainder were stationary. Each experiment involved 5000 replications and the critical value was taken as the 95th percentile. The average empirical critical value was then taken and is given in the table. T is the sample size in the replications.

We computed by Monte Carlo simulation the five percent empirical critical values for the MADF statistic for a four-variable system calibrated on the four actual real exchange rate data series (with unit roots imposed) for a sample size corresponding to that of our data set (288 observations), as well as for various other sample sizes for use in the Monte Carlo experiments discussed in Section 8, $T \in \{50, 75, 100, 200, 288, 300\}$.[18] The results are reported in panel (a) of Table 2. For the sample size of 288, we obtained a five percent critical value of 12.63. Applying the MADF test to the four real exchange rates themselves yielded a MADF statistic value of 15.88. We thus reject at the five percent level the joint null hypothesis that all of the series are generated by I(1) processes, implying that at least one of the real exchange rate series in the system is a realization of a stationary process.

In addition, we also examined a more general version of the MADF which allows a deterministic linear trend to be included in each of the N auxiliary regressions, and we term the resulting Wald statistic for the joint null hypothesis of a unit root in each of the series the MADFτ statistic. We calculated the five percent empirical critical values for the MADFτ test statistic for a four-variable system calibrated on our actual

[18] All Monte Carlo results discussed in this section were constructed using 5000 replications in each experiment, with identical random numbers across experiments. They were based on first-order systems calibrated on the actual real exchange rate data. Higher-order terms were insignificant in estimated linear autoregressions for all of the exchange rate series. Random error terms were drawn from a Gaussian distribution with a covariance matrix equal to the sample residual covariance matrix. The replications were initialized with the log real exchange rates set to zero (note that our log real exchange rate series were in fact constructed to have value zero in 1973M01), and artificial data were generated with a sample size of $100 + T$ and the first 100 data points discarded. For the MADF and MADFτ simulations, the system comprised four random walks. For the JLR simulations (see Section 5.2.2), the system comprised one, two, three, or four random walks, with the remainder mean-reverting AR(1) processes calibrated on the actual real exchange rate data.

real exchange rate data for the same set of sample sizes as before, and these are also reported in panel (a) of Table 2. For the sample size of 288, we obtained a five percent critical value of 16.37 for the MADFτ statistic, while applying this test to the four real exchange rate series themselves yielded a statistic value of 19.00, implying rejection of the joint null hypothesis of a unit root in each of the series at the five percent level, even allowing for deterministic trends.[19]

5.2.2. *JLR test.* The null hypothesis tested by the MADF and MADFτ test statistics—as with conventional panel unit root tests in general—is that all of the series under consideration are realizations of I(1) processes. Taylor and Sarno (1998) propose an alternative multivariate unit root test of the null hypothesis that at least one of the series is generated by an I(1) process. Essentially this is a special case of Johansen's (1988) likelihood ratio test for cointegration. Taylor and Sarno (1998) note that, if there are N cointegrating vectors among N series, then each of the series must be generated by an I(0) process and thereby form a cointegrating vector on its own.[20] They therefore suggest applying Johansen's (1988) likelihood ratio test statistic to test for N cointegrating vectors among N real exchange rate series, which is equivalent to testing the null hypothesis that at least one of the series is I(1), which is only violated when *all* of the series are I(0).[21] Taylor and Sarno (1998) term this the JLR test statistic.[22]

Taylor and Sarno (1998) show that the JLR statistic has a known limiting χ^2 distribution with one degree of freedom under the null hypothesis of at least one unit root process. Nevertheless, since it has been demonstrated that there may be substantial finite-sample bias of the Johansen statistic toward over-rejection of the null hypothesis (Cheung and Lai, 1993b), we also estimated its small-sample empirical distribution by Monte Carlo methods. The five percent critical values for the JLR statistic, again calculated by Monte Carlo simulations calibrated using the actual exchange rate data, are reported for a range of sample sizes in panel (b) of Table 2. As in Taylor and Sarno (1998), we find that the estimated critical values are in fact close to the corresponding asymptotic critical value from the $\chi^2(1)$ distribution for large sample sizes. Applying the JLR test statistic to our real exchange rate data, we obtained a test statistic of 4.33 which is larger than the relevant empirical critical value (for a sample size of 288) of 3.91 as well as the $\chi^2(1)$ asymptotic critical value of 3.84, enabling us to reject at the five percent significance level the null hypothesis that at least one of the series in the panel is a realization of a unit root process. Thus, the JLR test detects significant evidence of mean reversion in *each* of the real exchange rates.

[19] Note, however, that the deterministic trends were found to be both jointly and individually not statistically significantly different from zero at conventional nominal significance levels, suggesting therefore that the MADF statistic is the more appropriate in this case.

[20] Assuming, that is, the maintained hypothesis that each of the series is a realization of either an I(0) or an I(1) stochastic process, as is standard in this literature.

[21] Again subject to a maintained hypothesis that each of the series is a realization of either an I(0) or an I(1) process.

[22] Note that, for this null hypothesis, Johansen's two test statistics—the λ-max and λ-trace test statistics—coincide.

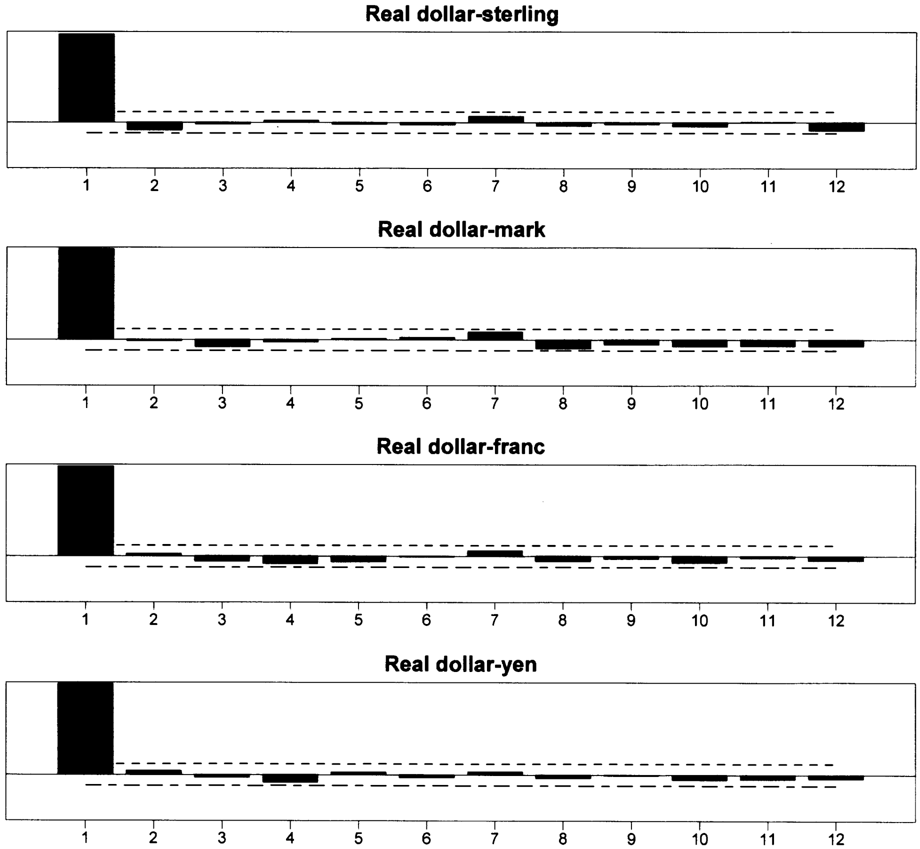

FIGURE 1

PARTIAL AUTOCORRELATION FUNCTIONS

We return in Section 8 to a discussion of the effect of nonlinearities on the MADF, MADFτ, and JLR test statistics.

6. NONLINEAR ESTIMATION RESULTS: RESOLVING THE FIRST PPP PUZZLE

The partial autocorrelation functions for each of the real exchange rates together with bands drawn at the two-sided five percent critical values (i.e., ± 1.96 divided by the square root of the sample size) are shown in Figure 1. They reveal that, in each case, only the first partial autocorrelation coefficient is significantly different from zero at the five percent level. Hence, for each of the real exchange rates, we set $p = 1$. Application of the Teräsvirta Rule then led to the choice of an ESTAR model with $p = d = 1$ for all four real exchange rates.[23] As a check on the robustness

[23] Detailed results of applying the Teräsvirta Rule are not reported to conserve space, but are available from the authors on request. In each case, there was no ambiguity in the choice of the model indicated.

of this choice, however, we also carried out a specification search for our nonlinear model as described in Section 3. First, setting $p = d = 1$ for each of the real exchange rates, we estimated models of the form

$$(8) \qquad q_t - \mu = \beta_1[q_{t-1} - \mu] + \beta_1^*[q_{t-1} - \mu][1 - \exp\{-\theta^2[q_{t-1} - \mu]^2\}] + \varepsilon_t$$

Experimentation with various starting values for the parameters yielded identical results, indicating the location of a global optimum. For each of the estimated ESTAR models, we could not reject, at the five percent significance level, the hypothesis of no remaining nonlinearity for values of d ranging from 2 to 12, on the basis of Lagrange multiplier tests (in Table 3 we report only the maximal value of the LM statistic testing for remaining ESTAR nonlinearity, $NLES_{\text{MAX}}$). Neither could we reject at the five percent level the hypothesis of no remaining nonlinearity of the LSTAR variety with values of the delay parameter in the range of 1 to 12 ($NLLS_{\text{MAX}}$ in Table 3). This procedure therefore suggested setting $d = 1$.

With $p = 1$, we also estimated the delay parameter, d, directly together with the other model parameters, by nonlinear least squares involving a grid search over values of d from 1 to 12. Using either single-equation or multivariate least squares, a value of d of unity was again implied for each of the real exchange rates.[24] It is significant that $d = 1$ is the least squares estimate because, as noted by Hansen (1997), since the parameter space for d is discrete, its least squares estimate is superconsistent and can be treated as known for the purposes of further inference.

Hence, the Teräsvirta Rule appears to be very robust in the present application and an ESTAR model with $p = d = 1$ was our preferred model for all of the series. The resulting ESTAR models, estimated jointly by multivariate nonlinear least squares,[25] are reported in Table 3.[26]

[24] It would also have been possible to extend this procedure to two dimensions and to search for the least squares estimates of both p and d. Given the strong evidence of no higher than first-order serial correlation in the real exchange rate series (Figure 1), however, this seemed unnecessary.

[25] The estimator employed is essentially the nonlinear seemingly unrelated regressions estimator, with iterations executed both on the parameters and the residual covariance matrix. Regularity conditions for the consistency and asymptotic normality of the nonlinear least squares estimator are discussed in the context of nonlinearity by Tjostheim (1986).

[26] As pointed out by an anonymous referee, the ESTAR model (8) (as well as the more general STAR model discussed in Section 3) is not the most general fomulation since the "location parameter" of the transition function is restricted to a single value equal to μ. While this seems a reasonable assumption to make in the present context, one might wish to allow for the possibility that the nonlinearity was in fact of the TAR variety. At the suggestion of the referee, therefore, we considered a related parameterization of the STAR model with the transition function

$$\Phi(\theta; \mu_1; \mu_2) = \left(1 + \exp\{-\theta^2(q_{t-d} - \mu_1)(q_{t-d} - \mu_2)\}\right)^{-1}$$

where $\mu_1 \leq \mu_2$. This transition function, advocated by Jansen and Teräsvirta (1996) and Teräsvirta (1998), is symmetric with respect to $(\mu_1 + \mu_2)/2$ and one might assume symmetry about μ, so that $\mu_1 = \mu - c$ and $\mu_2 = \mu + c$ where $c \geq 0$. An interesting feature of this transition function is that, as $\theta^2 \to \infty$, it approaches a double step function ($\Phi = 0$ for $\mu_1 \leq q_{t-d} - \mu \leq \mu_2$, $\Phi = 1$ elsewhere) and the corresponding STAR model then becomes a special case of a TAR model. Thus, this formulation essentially nests a TAR model of the type employed, for example, by Obstfeld and Taylor (1997) for testing for the law of one price with disaggregated goods price data. In estimation of a STAR model with the above transition function, however, we had severe problems in achieving

TABLE 3

NONLINEAR ESTIMATION RESULTS

(a) *Dollar–sterling*

$$\hat{q}_t = q_{t-1} - \left[1 - \exp\{\underset{\substack{(-2.771)\\[0.002]}}{-0.452}\{q_{t-1} + \underset{(4.274)}{0.149}\}^2\}\right][q_{t-1} + \underset{(4.274)}{0.149}]$$

$R^2 = 0.94$ $s = 0.033$ AR(1) = [0.63] AR(1 − 6) = [0.38]
$NLES_{\text{MAX}}$ = [0.23] $NLLS_{\text{MAX}}$ = [0.69]

(b) *Dollar–mark*

$$\hat{q}_t = q_{t-1} - \left[1 - \exp\{\underset{\substack{(-3.082)\\[0.001]}}{-0.257}q_{t-1}^2\}\right]q_{t-1}$$

$R^2 = 0.96$ $s = 0.034$ AR(1) = [0.75] AR(1 − 6) = [0.54]
$NLES_{\text{MAX}}$ = [0.31] $NLLS_{\text{MAX}}$ = [0.57]

(c) *Dollar–franc*

$$\hat{q}_t = q_{t-1} - \left[1 - \exp\{\underset{\substack{(-2.809)\\[0.003]}}{-0.294}\{q_{t-1} + \underset{(2.135)}{0.049}\}^2\}\right][q_{t-1} + \underset{(2.135)}{0.049}]$$

$R^2 = 0.96$ $s = 0.033$ AR(1) = [0.86] AR(1 − 6) = [0.87]
$NLES_{\text{MAX}}$ = [0.34] $NLLS_{\text{MAX}}$ = [0.51]

(d) *Dollar–yen*

$$\hat{q}_t = q_{t-1} - \left[1 - \exp\{\underset{\substack{(-2.652)\\[0.004]}}{-0.170}\{q_{t-1} + \underset{(11.208)}{0.516}\}^2\}\right][q_{t-1} + \underset{(11.208)}{0.516}]$$

$R^2 = 0.98$ $s = 0.034$ AR(1) = [0.75] AR(1 − 6) = [0.86]
$NLES_{\text{MAX}}$ = [0.44] $NLLS_{\text{MAX}}$ = [0.56]

(e) *System test statistics*

LR = [0.41] $W = \underset{[0.00]}{17.82}$ $SLR = \underset{[0.00]}{15.29}$

(f) *Covariance/correlation matrix of residuals*

dollar–sterling	0.001107	0.661529	0.681894	0.451354
dollar–mark	0.000757	0.001183	0.919155	0.526441
dollar–franc	0.000746	0.001039	0.001081	0.539532
dollar–yen	0.000510	0.000614	0.000602	0.001153

NOTES: A hat denotes the fitted value, R^2 denotes the coefficient of determination, and s is the standard error of the regression. AR(1) and AR(1 − 6) are Lagrange multiplier test statistics for first-order and up to sixth-order serial correlation in the residuals, respectively, constructed as in Eitrheim and Teräsvirta (1996). $NLES_{\text{MAX}}$ is the maximal Lagrange multiplier test statistic for no remaining ESTAR nonlinearity with delay in the range from 2 to 12 (Eitrheim and Teräsvirta, 1996). $NLLS_{\text{MAX}}$ is the maximal Lagrange multiplier test statistic for no remaining LSTAR nonlinearity with delay in the range from 1 to 12 (Eitrheim and Teräsvirta, 1996). LR is a likelihood ratio statistic for the nine parameter restrictions on the unrestricted ESTAR system. W is a joint Wald test statistic for the null hypothesis that the transition parameter is insignificantly different from zero in each of the estimated equations; SLR is a parametric bootstrap likelihood ratio test of the same null hypothesis constructed as suggested by Skalin (1998). Figures in parentheses below coefficient estimates denote the ratio of the estimated coefficient to the estimated standard error of the coefficient estimate. Figures given in square brackets denote marginal significance levels. For test statistics which are distributed as central χ^2 (the LR statistics) or F (the AR, $NLES_{\text{MAX}}$, and $NLLS_{\text{MAX}}$ statistics) under the null hypothesis, we report only the marginal significance level. Marginal significance levels for the "t-ratios" of the estimated transition parameters and for W and SLR were calculated by Monte Carlo methods as described in the text. In panel (f), we report the covariance/correlation matrix of the residuals (the upper triangle gives the correlations).

Table 3 in fact reports only the most parsimonious form of the estimated equations, since in no case could we reject at the five percent significance level the restrictions that $\beta_1 = -\beta_1^* = 1$ and, for the dollar-mark, $\mu = 0$ (see the likelihood ratio statistic, LR, in Table 3). These restrictions imply an equilibrium log-level of the real exchange rate of μ, in the neighborhood of which q_t is close to a random walk, becoming increasingly mean reverting with the absolute size of the deviation from equilibrium.

The residual diagnostic statistics are satisfactory in all cases (Eitrheim and Teräsvirta, 1996). The estimated standardized transition parameter in each case appears to be strongly significantly different from zero both on the basis of the individual "t-ratios" as well as on the basis of the very large Wald statistic, W, which tests for their joint significance. It should be clear on reflection, however, that the "t-ratios" for these parameters should be interpreted with caution since, under the null hypothesis $H_0 : \theta^2 = 0$, each of the real exchange rate series follows a unit root process. Hence, the presence of a unit root under the null hypothesis complicates the testing procedure analogously to the way in which the distribution of a Dickey–Fuller statistic cannot be assumed to be Student's t. Similarly, W cannot be assumed to be distributed as χ^2 under the null hypothesis. We therefore calculated the empirical marginal significance levels of these test statistics and of the Wald test statistic, W, by Monte Carlo methods assuming that the true data generating process for the logarithm of each of the four real exchange rate series was a random walk, with the parameters of the data generating process calibrated using the actual real exchange rate data over the sample period.[27] From these empirical marginal significance levels (reported in square brackets below the coefficient estimates in Table 3 and below the recorded value of W), we see that the estimated transition parameter is significantly different from zero at the one percent significance level in every case and the joint test that they are all zero has an empirical marginal significance level of zero to two decimal places. Since these tests may be construed as nonlinear univariate and multivariate unit root tests, the results indicate strong evidence of nonlinear mean reversion for each of the real exchange rates examined over the post-Bretton Woods period.

In addition, in order to provide corroborating evidence in favor of significant ESTAR nonlinearity in q and its nonlinear mean reversion to a stable equilibrium level, we also tested for the significance of θ^2 using Skalin's (1998) parametric bootstrap likelihood ratio test, generalized to a four-equation system. The resulting likelihood ratio statistic for the null hypothesis that $\theta^2 = 0$ for all four of the ESTAR

convergence both in single-equation and in multiple-equation estimation. This failure to converge suggests that smooth rather than discrete adjustment in regime is more appropriate, since it indicates difficulty in identifying the thresholds μ_1 and μ_2.

[27] The empirical significance levels were based on 5000 simulations of length 388, initialized at $q(1) = 0$ (calculations with $q(1)$ intialized at the estimated value of μ or at the actual value of the log real exchange rates in 1973M01 yielded virtually identical results), from which the first 100 data points were in each case discarded. At each replication a system of ESTAR equations identical in form to those reported in Table 3 was estimated. The percentage of replications for which a "t-ratio" for the estimated transition parameters greater in absolute value than that reported in Table 3 was obtained was then taken as the empirical significance level in each case. At the same time, we also calculated the percentage of replications in which the Wald test statistic, W, was greater than the recorded value of 17.82.

models, *SLR*, is very large, with a marginal significance level, similarly calculated by Monte Carlo Methods, of virtually zero, consistent with the Wald test statistic.[28]

Overall, the nonlinear estimation results are very interesting. We have uncovered statistically significant evidence of nonlinear mean reversion for each of the real exchange rates examined over the post-Bretton Woods period. The estimated models are in every case statistically well determined, provide good fits to the data, and pass a battery of diagnostic tests.[29,30]

7. ESTIMATING THE SPEED OF MEAN-REVERSION: RESOLVING THE SECOND PPP PUZZLE

While the estimated ESTAR models given in Table 3 impart some idea of the degree of mean reversion exhibited by the real exchange rates, a sensible way to gain a full insight into the mean-reverting properties of the estimated nonlinear models is through dynamic stochastic simulation. In particular, an analysis of the impulse response functions will allow the half-life of shocks to the real exchange rate models to be gauged and these can be compared to those previously reported in the literature in order to see if explicitly modeling nonlinearities helps resolve the purchasing power parity puzzle of very slow real exchange rate adjustment (Rogoff, 1996).

Thus, we examined the dynamic adjustment in response to shocks through impulse response functions which record the expected effect of a shock at time t on the system at time $t+j$. For a univariate linear model, the impulse response function is equivalent to a plot of the coefficients of the moving average representation (see, e.g., Hamilton, 1994: p. 318). Estimating the impulse response function for a nonlinear model raises special problems both of interpretation and of computation, however

[28] The evidence of nonlinearity implied by the tests involved in application of the Teräsvirta Rule, and by the univariate and multivariate tests for the statistical significance of the transition parameters ($H_0 : \theta^2 = 0$), was, in our view, strong enough to obviate the need to construct further tests such as those derived and discussed by Andrews and Ploberger (1994) and Hansen (1996).

[29] Qualitatively similar results are reported in Taylor and Peel (1998), using quarterly real dollar exchange rate data for the German mark, Japanese yen, and UK sterling deflated using the producer price deflator, and also in Taylor and Sarno (1999b), using data for German mark-based real exchange rates against the French franc, Dutch guilder, Belgian franc, and Italian lira for the post-Bretton Woods period as well as for the period of operation of the European Monetary System (EMS) since 1979. While EMS real exchange rates generally displayed faster adjustment speeds—in terms of higher estimated values of θ^2—a significant downward shift in the speed of mean-reversion after the EMS became operational in 1979 was detected (although no such shift was evident, for example, in the mark–sterling exchange rate). This finding accords with our economic intuition in that, although credibility of the EMS required mean reversion of the real exchange rate, its use as an anti-inflationary device meant that the brunt of adjustments to competitiveness had to be borne by movements in relative prices which are relatively sticky compared to nominal exchange rate movements under a free float (see, e.g., Artis and Taylor, 1988).

[30] We also used the estimated models to see if they were capable of generating the kind of dollar overvaluation seen during the 1980s—by some 42 percent, on our estimates, against the mark in February 1985, for example. In 10,000 simulated data series equal in length to the actual sample period, using the estimated ESTAR model for dollar–mark given in Table 3 with an initial value of zero, 15.9 percent of the simulated series reached an overvaluation of at least 42 percent during the sample period.

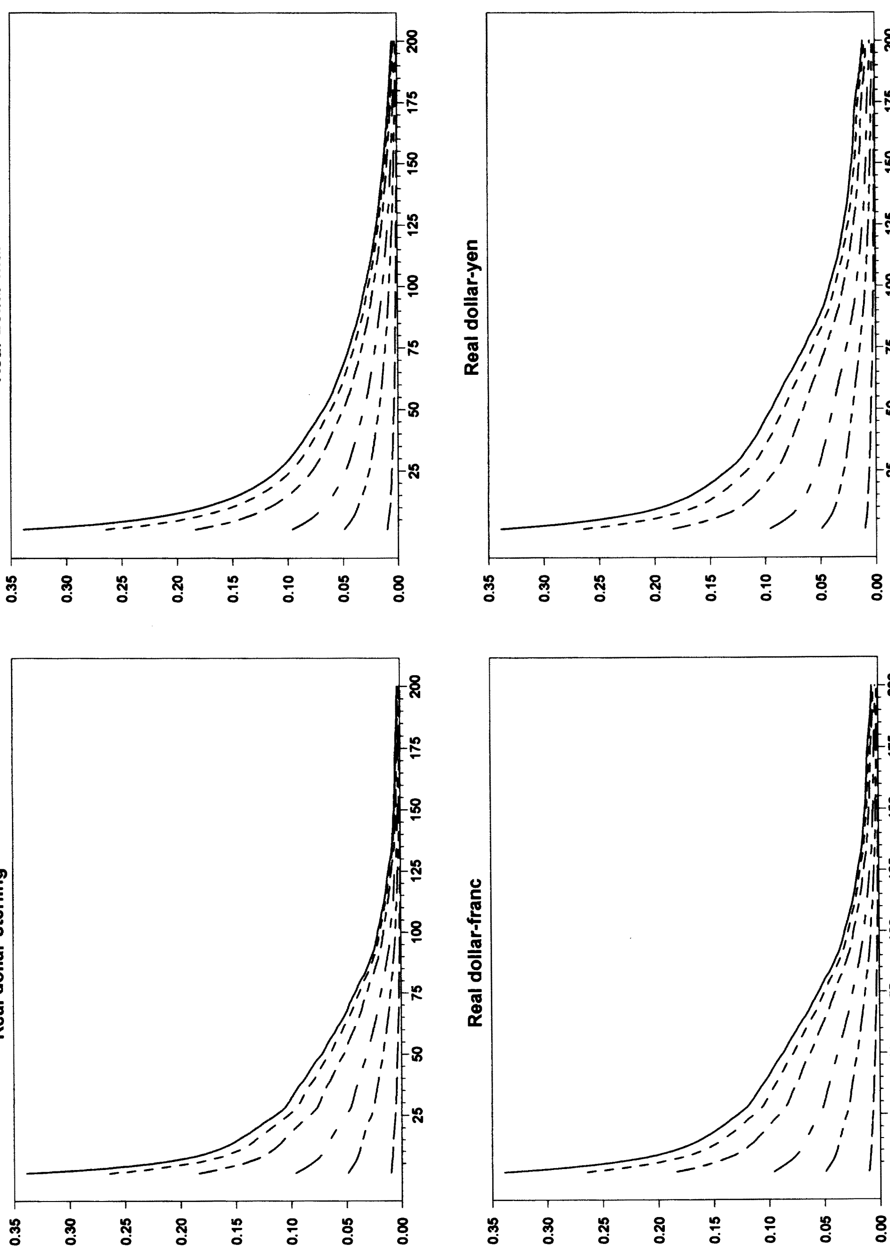

FIGURE 2

ESTIMATED IMPULSE RESPONSE FUNCTIONS

(Potter, 1991; Gallant et al., 1993; Koop et al., 1996). In particular, with nonlinear models, the shape of the impulse response function is not independent with respect to either the history of the system at the time the shock occurs, the size of the shock considered, or the distribution of future exogenous innovations. Exact estimates can only be produced—for a given shock size and initial condition—by multiple integration of the nonlinear function with respect to the distribution function each of the j future innovations, which is computationally impracticable for the long forecast horizons required in impulse response analysis. In this article, we calculate the impulse response functions, both conditional on average initial history and conditional on initial real exchange rate equilibrium, using the Monte Carlo integration method discussed by Gallant et al. (1993)—see the Appendix for further details.

The estimated impulse response functions are graphed in Figure 2 for each of the real exchange rates and six sizes of shock $k \in \{1, 5, 10, 20, 30, 40\}$, conditional on average initial history of the real exchange rates over the sample period. These graphs illustrate very clearly the nonlinear nature of the adjustment, with the impulse response functions for larger shocks decaying much faster than those for smaller shocks.

The estimated half-lives of the four real exchange rate models (Appendix) also illustrate the nonlinear nature of the estimated real exchange rate models, with larger shocks mean-reverting much faster than smaller shocks and shocks conditional on average history mean-reverting much faster than those conditional on initial equilibrium. The dollar–sterling and dollar–mark models show a marked degree of similarity in terms of the estimated half-lives. Conditional on average history, for real dollar–sterling and dollar–mark the models show quite fast mean-reversion, ranging from a half-life of under one year for the largest shocks of 40 percent to just under three years for very small shocks of 1 percent; for shocks of 5 to 10 percent, the half lives are just over two years. The dollar–franc displays slightly higher persistence, conditional on average history, with half-lives ranging from 13 months for a 40 percent shock to 40 months for a 1 percent shock while for dollar–yen the range is 14 to 42 months. For the other sizes of shocks considered, the half-lives for dollar–franc and dollar–yen are very close.

These results therefore seem to shed some light on what we have termed the second PPP puzzle (Rogoff, 1996). Only for small shocks occurring when the real exchange rate is near its equilibrium do our nonlinear models consistently yield half-lives in the range of three to five years, which Rogoff (1996) terms "glacial." For dollar–mark and dollar–sterling in particular, even small shocks of one to five percent have a half-life under three years, conditional on average history. For larger shocks, the speed of mean reversion is even faster.[31,32]

[31] Fitting linear autoregressive models to our data yields estimated speeds of real exchange rate adjustment in keeping with the literature—for dollar–sterling, for example, the half-life is just over 45 months. For the nonlinear ESTAR models, longer half-lives than those discussed above are recorded when the unrealistic assumption of initial exchange rate equilibrium is made—see the Appendix.

[32] Half-lives estimated using ESTAR models fitted to mark-based European real exchange rate series (Taylor and Sarno, 1999b) were generally slightly lower than those for dollar-based real exchange rates. This is unsurprising, given the proximity of the European markets involved and the fact that they are operating within a customs union, and accords with previous evidence on the

8. THE POWER OF UNIT ROOT TESTS WHEN EXCHANGE RATES ARE NONLINEARLY MEAN REVERTING: EXPLAINING THE LITERATURE

Given our discussion in Section 3 of the possibility of nonrejection of the unit root hypothesis when in fact the true process is nonlinearly mean reverting, it seems worthwhile to investigate the frequency with which the hypothesis of a unit root can be rejected using standard test procedures when, under the null hypothesis, the data generating process is calibrated according to our estimated ESTAR models. This may help us understand why much previous research has resulted in nonrejection (at conventional test sizes) of the unit root hypothesis for real exchange rates over the floating rate period using standard univariate linear unit root tests. It may also shed some light on the findings of the more recent literature applying panel and other multivariate unit root tests to real exchange rate data.

8.1. *The Power of Univariate Unit Root Tests.* We executed a number of Monte Carlo experiments based on an artificial data generating process identical to the four estimated ESTAR models reported in Table 3, with independent and identically distributed Gaussian innovations with a covariance matrix as given in panel (f) of Table 3.[33] Initializing the artificial series at zero, we generated 5000 samples of 400 observations and discarded the first 100, leaving 5000 samples of 300 observations. For each generated sample of observations we then calculated an optimized augmented Dickey–Fuller statistic in which from 1 to 12 lags of first differences were added into the auxiliary regression until a Lagrange multiplier test for up to sixth-order serial correlation was insignificant at the five percent level.[34] The ADF statistic was then calculated and rejection frequencies over the whole 5000 samples were computed, using five percent critical values calculated from the response surface estimates of MacKinnon (1991). We then repeated the calculation of the rejection frequencies using only 100, 200, and 288 data points.

The results of these Monte Carlo investigations are given in Table 4 and reveal a distinct lack of power of univariate ADF tests to reject a false null hypothesis of unit root behavior when the process is in fact nonlinearly mean reverting (panel (a)). For a sample size of 200, corresponding to the number of monthly data points available for the recent period of floating exchange rates (i.e., since 1973) at the end of the 1980s, the rejection frequencies range from about 28 to 38 percent.[35] That is to say, there would be approximately a 60 to 70 percent probability of failing to reject the hypothesis of unit root behavior of real exchange rates if in fact they were generated

mean-reverting properties of European real exchange rates (e.g., Canzoneri et al., 1999; Cheung and Lai, 1998).

[33] Statistical tests revealed no evidence of non-normality or heteroskedasticity in the estimated residual series.

[34] A constant term was included in the auxiliary regressions. Qualitatively identical results were obtained when a deterministic trend was also included.

[35] In fact, many of the studies published around the end of the 1980s which applied univariate unit root tests to real exchange rates used data sets ending a few years earlier, in which case the power would be even lower.

TABLE 4

POWER FUNCTIONS OF UNIT ROOT TESTS WHEN THE TRUE MODEL IS ESTAR

(a) *Univariate ADF test*	$T = 100$	$T = 200$	$T = 288$	$T = 300$
UK	14.20	38.04	49.18	52.16
Germany	9.26	34.32	43.80	50.02
France	7.25	28.74	40.24	43.12
Japan	10.78	37.14	48.68	51.40

(b) *MADF and MADFτ tests*	$T = 100$	$T = 200$	$T = 288$	$T = 300$
MADF				
*DGP*1	99.80	100.00	100.00	100.00
*DGP*2	95.72	100.00	100.00	100.00
*DGP*3	90.10	100.00	100.00	100.00
*DGP*4	89.16	99.40	100.00	100.00
MADFτ				
*DGP*1	97.22	100.00	100.00	100.00
*DGP*2	92.80	100.00	100.00	100.00
*DGP*3	88.16	100.00	100.00	100.00
*DGP*4	85.24	98.20	100.00	100.00

(c) *JLR test*	$T = 100$	$T = 200$	$T = 288$	$T = 300$
	33.10	81.26	100.00	100.00

NOTES: In all cases rejection frequencies are at a nominal significance level of five percent and are calculated on the basis of 5000 replications, with identical random numbers used across simulations. The precise details of the simulations are given in the text. T is the sample size. *Panel* (a): Critical values were calculated from the response surface estimates of MacKinnon (1991). The data generating process is, for each country considered, an ESTAR based on the estimates reported in Table 3. *Panel* (b): Critical values are as given in panel (a) of Table 2. $DGPi(i = 1, \dots, 4)$ denotes a data generating process comprising $(5 - i)$ ESTAR processes, with the remaining $(i - 1)$ processes following a random walk. *Panel* (c): Critical values are as given in panel (b) of Table 2. The data generating process comprises four ESTAR processes.

by nonlinearly mean-reverting ESTAR models of the kind we have estimated. This does not mean that we should expect to find 30 percent of empirical studies rejecting unit roots in real exchange rates since, in actual fact, we only have one set of real exchange rate data for the recent float. Rather, it means that, if real exchange rates were in fact generated by processes close to our estimated ESTAR models, there is a 70 percent probability that our data set would *never* allow us to reject the unit root hypothesis using standard univariate unit root tests. The presence of nonlinearities in real exchange rate adjustment may therefore explain why a number of papers published at the end of the 1980s (e.g., Taylor, 1988; Mark, 1990) were unable to reject the unit root hypothesis using real exchange rate data for the post-Bretton Woods float.

The power of the optimized ADF test hardly improves, moreover, when we extend the sample up to 288 (the size of our data set) or beyond, to 300. The probability of rejection of the false null hypothesis in this case rises to at most around 50 percent.[36]

8.2. *Power of Multivariate Unit Root Tests.*

8.2.1. *MADF and MADFτ tests.* We also investigated the power of panel unit root tests in the presence of nonlinear mean reversion, using our MADF and MADFτ tests as representative of this class of tests. In this case, since the null hypothesis is that each of the four processes is I(1), it can be violated by nonlinear mean reversion when one, two, three, or all four of the processes are nonlinearly mean reverting. Accordingly, we calculated the power functions for these tests using separate data generating processes. The first data generating process, DGP1, had four ESTAR processes, identical to those reported in Table 3; the second, DGP2, had three, identical to those estimated for dollar–sterling, dollar–mark, and dollar–franc, and a fourth random walk process calibrated on the dollar–yen data; the third, DGP3, had two, identical to those estimated for dollar–sterling and dollar–mark, and two random walk processes calibrated on the dollar–franc and dollar–yen data; and the fourth, DGP4, had one ESTAR process identical to that estimated for dollar–sterling, and three random walk processes calibrated on the other exchange rates. The method of calculating the rejection frequencies was otherwise similar to that discussed in the previous subsection.

The results are given in panel (b) of Table 4 and do indeed reveal the MADF and MADFτ statistics to be very powerful even in the face of nonlinear mean reversion. However, as in Taylor and Sarno (1998), the pitfall in panel unit root tests of this class is also evident from Table 4: the rejection frequencies for the MADF and MADFτ statistics are virtually 100 percent for sample sizes of 200 or more even when as few as one of the processes is (nonlinearly) mean reverting and three of the processes are following random walks.

8.2.2. *JLR test.* Finally, we examined the power of the JLR test to reject the false null hypothesis that at least one of the processes is I(1) when in fact they are all nonlinearly mean-reverting, using a data generating process identical to our four estimated ESTAR models, and calculating the rejection frequencies as previously. The results are given in panel (c) of Table 4 and show that the JLR test remains very powerful in the face of nonlinear mean reversion. While the JLR test is less powerful than the MADF or MADFτ tests for sample sizes of 200 or less, for a sample size corresponding to that of our actual data set (288), the rejection frequency is 100 percent, indicating that the JLR statistic may be usefully applied even in the presence

[36] The results reported in this subsection are consistent with and reinforce the Monte Carlo evidence reported by Skalin and Teräsvirta (1998) on the low power of univariate unit root tests in the presence of nonlinear mean reversion, although Skalin and Teräsvirta's (1998) analysis is based on an application of LSTAR models, since they are primarily concerned with modelling asymmetries in movements in unemployment. This suggests that our Monte Carlo results are not particularly subject to the problem of specificity (Hendry, 1984).

of nonlinear mean reversion, while avoiding the pitfall in standard panel unit root which we have noted.[37,38]

9. CONCLUSION

Our empirical results provide strong confirmation that four major real bilateral dollar exchange rates are well characterized by nonlinearly mean reverting processes over the floating rate period since 1973.[39] The crucial estimated parameters, the transition parameters, were of the correct signs and plausible magnitudes and were shown to be strongly statistically significantly different from zero, allowing for a unit root process under the null hypothesis and calculating their empirical significance levels by Monte Carlo methods. The estimated models imply an equilibrium level of the real exchange rate in the neighborhood of which the behavior of the log-level of the real exchange rate is close to a random walk, becoming increasingly mean reverting with the absolute size of the deviation from equilibrium. This is consistent with a number of recent theoretical contributions on the nature of real exchange rate dynamics in the presence of international arbitrage costs.

We estimated the impulse response functions corresponding to our estimated nonlinear real exchange rate models by Monte Carlo integration. Because of the nonlinearity, the half-lives of shocks to the real exchange rates vary both with the size of the shock and with the initial conditions. By taking account of statistically significant nonlinearities, we find the speed of real exchange rate adjustment to be typically much faster than the apparently "glacial" speeds of real exchange rate adjustment recorded hitherto (Rogoff, 1996).

In a number of Monte Carlo studies calibrated on the estimated nonlinear models, we also demonstrated the very low power of standard univariate unit root tests to reject a false null hypothesis of unit root behavior when the true model is nonlinearly mean-reverting, thereby suggesting an explanation for the difficulty researchers have encountered in rejecting the linear unit root hypothesis at conventional significance levels for major real exchange rates over the recent floating rate period. Panel unit root tests, however, displayed much higher power in their rejection of the false null hypothesis against an alternative of nonlinear mean-reversion, in keeping with the recent literature. Our results therefore encompass previous empirical work in this area.[40]

While our results aid our understanding of real exchange rate behavior over the recent float, we view them as a tentatively adequate characterization of the data which

[37] See Balke and Fomby (1997) for a related Monte Carlo study of conventional cointegration testing procedures in the context of threshold cointegration.

[38] Calibrating the Monte Carlo experiments using other real exchange rate data (Taylor and Peel, 1998; Taylor and Sarno, 1999a, 1999b) yielded qualitatively similar results, again suggesting that our results do not suffer particularly from specificity (Hendry, 1984).

[39] Similar results also hold among intra-European German mark-based real exchange rates (Taylor and Sarno, 1999b).

[40] In our fitted ESTAR models, the real exchange rate will be closer to a unit root process the closer it is to its long-run equilibrium. Somewhat paradoxically, therefore, failure to reject a unit root may indicate that the real exchange rate has on average been relatively close to equilibrium, rather than implying that no such long-run equilibrium exists.

appears to be superior to linear exchange rate modelling in a number of respects, but which nevertheless may be capable of improvement. In particular, we may gain further insights into the adjustment process by developing nonlinear error correcting systems of equations involving prices, exchange rates, and other macroeconomic variables. These challenges remain on the agenda for future research.

APPENDIX: ESTIMATION OF HALF-LIVES OF REAL EXCHANGE RATE SHOCKS

This appendix describes the procedure employed to calculate the impulse response functions discussed in Section 7 and graphed in Figure 2.

We employ Monte Carlo methods to forecast a path for q_{t+j} (Granger and Teräsvirta, 1993) given h_t and with and without a shock of size s_t at time t. In our first estimation of the impulse response functions we condition on initial equilibrium by setting the initial lagged values of q equal to the estimated equilibrium level $\hat{\mu}$. A base run is then constructed by simulating the four estimated equations reported in Table 3 for 200 periods, with drawings from a vector of Gaussian innovations with covariance matrix identical to that given in panel (f) of Table 3. This procedure is then repeated with identical random numbers but with an additional additive shock at time t. The sequence representing the difference between the second simulated path of q and the base run at each horizon is then stored. The whole procedure is then repeated 5000 times and the average of the 5000 realizations of the sequence of differences taken as the estimated impulse response function. Since we use a large number of simulations, by the Law of Large Numbers this procedure should produce results virtually identical to that which would result from calculating the exact response functions analytically by multiple integration (Gallant et al., 1993).

This procedure is modified as follows in order to produce an estimate of the impulse response function conditional on the average history of each of the real exchange rates. Starting at the first data point, q_{t-1} is set equal to $\{| q(1973\text{M}01) - \hat{\mu} | + \hat{\mu}\}$. If $q(1973\text{M}01) - \hat{\mu}$ is positive, this is just $q(1973\text{M}01)$ itself; if $q(1973\text{M}01) - \hat{\mu}$ is negative, then $\{| q(1973\text{M}01) - \hat{\mu} | + \hat{\mu}\}$ is the value which is an equal absolute distance above the estimated equilibrium value $\hat{\mu}$. This transformation is necessary because we consider only positive shocks and it is innocuous because of the symmetric nature of ESTAR adjustment below and above equilibrium. Two hundred simulations of length 200, with and without a positive shock of size s_t at time t, are then generated using the estimated ESTAR model, and realizations of the differences between the two simulated paths are calculated and stored as before. We then move up one data point (hence setting $t - 1 = 1973\text{M}02$) and repeat this procedure. Once this has been done for every data point in the sample up to 1996M12, an average over all of the simulated sequences of differences in the paths of the real exchange rates with and without the shock at time t is taken as the estimated impulse response function conditional on the average history of the given exchange rate and for a given shock size. In all, this procedure requires $288 \times 200 = 57{,}600$ simulations for each of the real exchange rates and each size of shock.

As noted in Section 7, for linear time series models the size of shock used to trace out an impulse response function is not of particular interest since it serves only as a scale factor, but it is of crucial importance in the nonlinear case. In the present application we are particularly concerned with the effect of shocks to the level of the real exchange rate. Given a particular value of the log real exchange rate at time t, q_t—whether this be the historical value or the estimated equilibrium level—a shock of k percent to the level of the real exchange rate involves augmenting q_t additively by $s_t = \log(1 + k/100)$.[41]

This raises a problem, however, in the calculation of the half-lives since although the natural measure might be the discrete number of months taken until the shock to the level of the real exchange rate has dissipated by a half—i.e., when the impulse response function falls below $\log(1 + k/200)$—this would make comparison with research on linear time series models of real exchange rates difficult. Accordingly, although we define a k percent shock to the real rate as equivalent to adding $\log(1 + k/100)$ to q_t, we calculate the half-life as the discrete number of months taken for the impulse response function to fall below $0.5 \log(1 + k/100)$, facilitating a comparison of our results with half-lives estimated in previous studies. We considered six different sizes of percentage shock to the level of the real exchange rate, $k \in \{1, 5, 10, 20, 30, 40\}$. This allows us to compare and contrast the persistence of very large and very small shocks.

Implementation of these methods on our data produced the following estimated half-lives conditional on average initial history

Shock (%)	40	30	20	10	5	1
UK	10	20	22	26	29	32
Germany	11	14	18	25	29	33
France	13	18	24	33	38	40
Japan	14	18	24	32	38	42

and the following estimated half-lives conditional on initial exchange rate equilibrium

Shock (%)	40	30	20	10	5	1
UK	22	30	34	40	42	43
Germany	33	30	48	54	55	56
France	30	37	45	51	53	54
Japan	44	52	60	66	68	69

These estimated half-lives form the basis of our discussion in Section 7.

REFERENCES

ABUAF, N., AND P. JORION, "Purchasing Power Parity in the Long Run," *Journal of Finance* 45 (1990), 157–74.

ANDREWS, D. W. K., AND W. PLOBERGER, "Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative," *Econometrica* 62 (1994), 1383–1414.

[41] For small k, $\log(1 + k/100)$ is of course approximately equal to $k/100$. This approximation is not, however, good for the larger shocks considered in this article.

ARTIS, M. J., AND M. P. TAYLOR, "Exchange Rates, Interest Rates, Capital Controls and the European Monetary System: Assessing the Track Record," in F. Giavazzi, S. Micossi, and M. Miller, eds., *The European Monetary System* (Cambridge: Cambridge University Press, 1988).

BALKE, N. S., AND T. B. FOMBY, "Threshold Cointegration," *International Economic Review* 38 (1997), 627–45.

BAXTER, M., AND A. C. STOCKMAN, "Business Cycles and the Exchange Rate System," *Journal of Monetary Economics* 23 (1989), 377–400.

BENNINGA, S., AND A. A. PROTOPAPADAKIS, "The Equilibrium Pricing of Exchange Rates and Assets When Trade Takes Time," *Journal of International Economics* 7 (1988), 129–49.

BERTOLA, G., AND R. J. CABALLERO, "Kinked Adjustment Costs and Aggregate Dynamics," in O. J. Blanchard, and S. Fischer, eds., *NBER Macroeconomics Annual* (Cambridge, MA: MIT Press, 1990).

CANZONERI, M. B., R. E. CUMBY, AND B. DIBA, "Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries," *Journal of International Economics* 47 (1999), 245–66.

CASSEL, G., "Abnormal Deviations in International Exchanges," *Economic Journal* 28 (1918), 413–15.

———, *Money and Foreign Exchange after 1914* (London: Constable, 1922).

CHEUNG, Y. W., AND K. S. LAI, "A Fractional Cointegration Analysis of Purchasing Power Parity," *Journal of Business Economics and Statistics* 11 (1993a), 103–12.

——— AND ———, "Finite-Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration," *Oxford Bulletin of Economics and Statistics* 55 (1993b), 313–28.

——— AND ———, "Parity Reversion in Real Exchange Rates During the Post-Bretton Woods Period," *Journal of International Money and Finance* 17 (1998), 597–614.

COLEMAN, A. M. G., "Arbitrage, Storage and the Law of One Price: New Theory for the Time Series Analysis of an Old Problem," discussion paper, Department of Economics, Princeton University, 1995.

DAVUTYAN, N., AND J. PIPPENGER, "Testing Purchasing Power Parity: Some Evidence of the Effects of Transactions Costs," *Econometric Reviews* 9 (1990), 211–40.

DIEBOLD, F. X., S. HUSTED, AND M. RUSH, "Real Exchange Rates under the Gold Standard," *Journal of Political Economy* 99 (1991), 1252–71.

———, J.-H. LEE, and G. C. WEINBACH, "Regime Switching with Time-Varying Transition Probabilities," in C. P. Hargreaves, ed., *Nonstationary Time Series Analysis and Cointegration* (Oxford and New York: Oxford University Press, 1994).

DIXIT, A. K., "Hysteresis, Import Penetration and Exchange Rate Pass-Through," *Quarterly Journal of Economics* 104 (1989), 205–28.

DORNBUSCH, R., *Open Economy Macroeconomics* (New York: Basic Books, 1980).

DUMAS, B., "Dynamic Equilibrium and the Real Exchange Rate in Spatially Separated World," *Review of Financial Studies* 5 (1992), 153–80.

———, "Partial Equilibrium Versus General Equilibrium Models of the International Capital Market," in F. Van Der Ploeg, ed., *The Handbook of International Macroeconomics* (Oxford: Blackwell, 1994), Chap. 10.

EITRHEIM, O., AND T. TERÄSVIRTA, "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics* 74 (1996), 59–75.

ELWOOD, S. K., "Is the Persistence of Shocks to Output Asymmetric?," *Journal of Monetary Economics* 41 (1998), 411–26.

ENGEL, C., "Accounting for U.S. Real Exchange Rate Changes," *Journal of Political Economy* 107 (1999), 507–38.

ENGLE, R. F., AND C. W. J. GRANGER, "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica* 55 (1987), 251–76.

FLOOD, R. P., AND M. P. TAYLOR, "Exchange Rate Economics: What's Wrong with the Conventional Macro Approach?," in J. A. Frankel, G. Galli, and A. Giovannini, eds., *The Microstructure of Foreign Exchange Markets* (Chicago: Chicago University Press, 1996).

FRANKEL, J. A., "International Capital Mobility and Crowding out in the US Economy: Imperfect Integration of Financial Markets or Goods Markets?," in R. W. Hafer, ed., *How Open is the US Economy?* (Lexington: Lexington Books, 1986).

———, "Zen and the Art of Modern Macroeconomics: The Search for Perfect Nothingness," in W. Haraf, and T. Willett, eds., *Monetary Policy for a Volatile Global Economy* (Washington, DC: American Enterprise Institute, 1990).

——— AND A. K. ROSE, "A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries," *Journal of International Economics* 40 (1996), 209–24.

FROOT, K. A., AND K. ROGOFF, "Perspectives on PPP and Long-run Real Exchange Rates," in K. Rogoff, and G. Grossman, eds., *Handbook of International Economics* (Amsterdam: North-Holland, 1995).

FULLER, W. A., *Introduction to Statistical Time Series* (New York: Wiley, 1976).

GALLANT, A. R., P. E. ROSSI, and G. TAUCHEN, "Nonlinear Dynamic Structures," *Econometrica* 61 (1993), 871–908.

GRANGER, C. W. J., AND T. TERÄSVIRTA, *Modelling Nonlinear Economic Relationships* (Oxford: Oxford University Press, 1993).

HAMILTON, J. D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57 (1989), 357–84.

———, *Time Series Analysis* (Princeton NJ: Princeton University Press, 1994).

HANSEN, B. E., "Inference when a Nuisance Parameter Is not Identified under the Null Hypothesis," *Econometrica* 64 (1996), 413–30.

———, "Inference in TAR Models," *Studies in Nonlinear Dynamics and Econometrics* 2 (1997), 1–14.

HECKSCHER, E. F., "Växelkursens Grundval vid Pappersmyntfot," *Ekono-misk Tidskrift* 18 (1916), 309–12.

HEGWOOD, N. D., AND D. H. PAPELL, "Quasi Purchasing Power Parity," *International Journal of Finance and Economics* 3 (1998), 279–89.

HENDRY, D. F., "Monte Carlo Experimentation in Econometrics," in Z. Griliches and M. D. Intrilligator, eds., *Handbook of Econometrics* (Amsterdam: North-Holland, 1984).

JANSEN, E. S., AND T. TERÄSVIRTA, "Testing Parameter Constancy and Super Exogeneity in Econometric Equations," *Oxford Bulletin of Economics and Statistics* 58 (1996), 735–63.

JOHANSEN, S., "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control* 12 (1988), 231–54.

KOOP, G., M. H. PESARAN, AND S. POTTER, "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics* 74 (1996), 119–47.

KRUGMAN, P. R., *Exchange Rate Instability* (Cambridge, MA: MIT Press, 1989).

LOTHIAN, J. R., AND M. P. TAYLOR, "Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries," *Journal of Political Economy* 104 (1996), 488–510.

——— AND ———, "Real Exchange Rate Behavior: The Problem of Power and Sample Size," *Journal of International Money and Finance* 16 (1997), 945–54.

——— AND ———, "Purchasing Power Parity Over Two Centuries: Strengthening the Case for Real Exchange Rate Stability," *Journal of International Money and Finance* 19 (2000), 759–64.

LUUKKONEN, R., P. SAIKKONEN, AND T. TERÄSVIRTA, "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika* 75 (1988), 491–99.

MACKINNON, J., "Critical Values for Cointegration Tests," in R. F. Engle, and C. W. J. Granger, eds., *Long-Run Economic Relationships* (Oxford: Oxford University Press, 1991).

MARK, N., "Real and Nominal Exchange Rates in the Long Run: An Empirical Investigation," *Journal of International Economics* 28 (1990), 115–36.

MICHAEL, P., A. R. NOBAY, AND D. A. PEEL, "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation," *Journal of Political Economy* 105 (1997), 862–79.

OBSTFELD, M., AND K. ROGOFF, "Exchange Rate Dynamics Redux," *Journal of Political Economy* 103 (1995), 624–60.

——— AND ———, *Foundations of International Macroeconomics* (Cambridge, MA: MIT Press, 1996).

——— AND A. M. TAYLOR, "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited," *Journal of the Japanese and International Economies* 11 (1997), 441–79.

O'CONNELL, P. G. J., "Perspectives on Purchasing Power Parity," Ph.D. dissertation, Harvard University, 1997.

——, "The Overvaluation of Purchasing Power Parity," *Journal of International Economics* 44 (1998), 1–19.

—— AND S.-J. WEI, "The Bigger They Are, The Harder They Fall: How Price Differences Between US Cities are Arbitraged," discussion paper, Department of Economics, Harvard University (paper presented at the CEPR-PIES Conference on "Market integration and Real Exchange Rates," Georgetown University, Washington DC, 2–3 May 1997).

OFFICER, L. H., *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance* (Greenwich, CT: JAI Press, 1982).

OHANIAN, L. E., AND A. C. STOCKMAN, "Arbitrage Costs and Exchange Rates," discussion paper, Department of Economics, University of Rochester (paper presented at the CEPR-PIES Conference on "Market integration and Real Exchange Rates," Georgetown University, Washington DC, 2–3 May 1997).

PAPELL, D. H., "Searching for Stationarity: Purchasing Power Parity Under the Current Float," *Journal of International Economics* 43 (1998), 313–32.

PESARAN, M. H., AND S. M. POTTER, "A Floor and Ceiling Model of US Output," *Journal of Economic Dynamics and Control* 21 (1997), 661–96.

POTTER, S. M., "Nonlinear Impulse Response Functions," discussion paper, University of California, Los Angeles, 1991.

ROGOFF, K., "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34 (1996), 647–68.

SERCU, P., R. UPPAL, AND C. VAN HULLE, "The Exchange Rate in the Presence of Transactions Costs: Implications for Tests of Purchasing Power Parity," *Journal of Finance* 50 (1995), 1309–19.

SKALIN, J., "Testing Linearity against Smooth Transition Autoregression Using a Parametric Bootstrap," Working Paper Series in Economics and Finance 276, Department of Economic Statistics, Stockholm School of Economics, 1998.

—— AND T. TERÄSVIRTA, "Modelling Asymmetries and Moving Equilibria in Unemployment Rates," Working Paper Series in Economics and Finance 262, Department of Economic Statistics, Stockholm School of Economics, 1998.

TAYLOR, M. P., "An Empirical Examination of Long-Run Purchasing Power Parity Using Cointegration Techniques," *Applied Economics* 20 (1988), 1369–81.

——, "The Economics of Exchange Rates," *Journal of Economic Literature* 33 (1995), 13–47.

——, AND P. C. MCMAHON, "Long-Run Purchasing Parity in the 1920s," *European Economic Review* 32 (1988), 179–97.

—— AND D. A. PEEL, "Nonlinear Mean-Reversion in Real Exchange Rates," mimeo, Department of Economics, University of Oxford, 1998.

—— AND L. SARNO, "The Behavior of Real Exchange Rates During the Post-Bretton Woods Period," *Journal of International Economics* 46 (1998), 281–312.

—— AND ——, "Linear and Nonlinear Mean Reversion in Real Exchange Rates: Reconciling the Empirical Evidence," mimeo, Department of Economics, University of Oxford, 1999a.

—— AND ——, "Nonlinearities in European Real Exchange Rate Adjustment," mimeo, Department of Economics, University of Oxford, 1999b.

TERÄSVIRTA, T., "Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association* 89 (1994), 208–18.

——, "Modelling Economic Relationships with Smooth Transition Regressions," in D. E. A. Giles, and A. Ullah, eds., *Handbook of Applied Economic Statistics* (New York: Dekker, 1998).

TJOSTHEIM, D., "Estimation in Nonlinear Time Series Models," *Stochastic Processes and Their Applications* 21 (1986), 251–73.

TONG, H., *Nonlinear Time Series: A Dynamical System Approach* (Oxford: Clarendon Press, 1990).

WILLIAMS, J. C., AND B. D. WRIGHT, *Storage and Commodity Markets* (Cambridge: Cambridge University Press, 1991).

WU, Y., "Are Real Exchange Rates Non-Stationary?: Evidence from a Panel-Data Test," *Journal of Money, Credit and Banking* 28 (1996), 54–63.